

Result No.	Score	Query		DB	ID	Description
		Match	Length			
c	1	952.8	61.6	1164	1	Q03159 Complete gene sequ
	2	946.2	61.2	1164	1	S. griseofuscus gl
	3	931.8	60.3	1167	1	DNA sequence encod
	4	903	58.4	2010	1	Glucose isomerase
	5	599.6	38.8	1567	1	Sequence of the ri
	6	598.8	38.7	1185	1	Nucleotide sequenc
	7	598.4	38.7	1182	1	Complete gene sequ
	8	481.8	31.2	1805	1	Thermus aquaticus
	9	100.8	6.5	114955	1	Human adenosine Al
	10	97	6.3	1821	1	S. lividans protea
	11	97	6.3	1821	1	Clone P5-6 encodin
	12	93	6.0	24379	1	Streptomyces freno
	13	93	6.0	24379	1	Streptomyces roseo
	14	85.2	5.5	44377	1	Streptomyces roseo
	15	85.2	5.5	44377	1	Platenolide syntha
	16	83.8	5.4	114955	1	Platenolide syntha
	17	82.8	5.4	4257	1	Human adenosine Al
	18	82.8	5.4	4257	1	Infected cell prot
	19	80.6	5.2	1575	1	The nucleotide seq
	20	78.6	5.1	1308	1	Actinomycete phosp
	21	78.6	5.1	1308	1	S. lividans protea
	22	78.6	5.1	2634	1	Tripeptidyl aminop
	23	78.4	5.1	8169	1	Microbispora therm
	24	77.4	5.0	30001	1	Actinomadura hibi
	25	77.4	5.0	30001	1	Total DNA sequenc
	26	77	5.0	833	1	S. aureofaciens DNA
	27	77	5.0	12001	1	snaB gene encoding
	28	76.6	5.0	15872	1	HSV L/ST region. H
	29	76.4	4.9	1998	1	Streptomyces venez
	30	76.4	4.9	5392	1	Sequence encoding
	31	76.4	4.9	53789	1	Sequence compris
	32	75.2	4.9	12588	1	Amycolatopsis medi
	33	74.6	4.8	15079	1	Sequence encoding
	34	73.6	4.8	2522	1	S. clavuligerus cl
					Hyperthermostable	





```
Db 901 GCTCGGCGCGGCTGCATGCGCAACTACCTGATCTCTAAGAGAGTGCGCGCCCTTC 960
Qy 1220 qctgcgacccgaggtccagagagccctcgctgcgcccgcctcgacagctcgccag 1279
Db 961 CCGCGCGACCCGAGGTGAGAGAGCGGTGCGCGGTCCCGTCTGAGAGAGTGCCCGG 1020
Qy 1280 cccacccgcccgcagcgcctcgagcgcctcgagcgcctggtgcccagccagcagcgcgtacgagacttc 1339
Db 1021 CCACGCGCGCGAGCGGTCTGCACCCCTGCTCGAGACAGCGGTCCGCTTCGAGAGTTC 1080
Qy 1340 gacgtgagcggcgccg--cggcgatgacctcgagcgcctcgagcgcctcgacagctcgccatg 1396
Db 1081 GACGTGAGCGCGCGCGCGCGGTGCGGTTCGAGCGCGCTGAGCGCTGAGTGCATG 1140
Qy 1397 gaccacctgctggcgccgcccgtgga 1423
Db 1141 GACCACCTGCTGGCGCGCGCGGCTGA 1167

RESULT 4
ID N90362 standard; DNA; 2010 BP.
AC N90362;
DF 1-NOV-1989 (first entry)
DE Glucose isomerase gene from Streptomyces spp
KW Glucose isomerase; Streptomyces; DNA.
OS Streptomyces
FH Key Location/Qualifiers
FT cds 566..1847
FT signal_peptide 566..675
FT /tag= a
FT /tag= b
PN J01137979-A.
PD 24-MAY-1989.
PF 24-NOV-1987; 295739.
PR 24-NOV-1987; JP-295739.
PA (NORQ) Norinaho.
DR WPI; 89-198224/27.
DR P-PSDB; P90423.
PT Novel glucose isomerase gene
PT - exists in chromosome of Streptomyces and is
PT contained in recombinant etc.
PS Claim; fig 1; pp.
CC Glucose isomerase gene from Streptomyces spp. (see P90423). Used to
CC study this protein's structure by recombinant techniques, and for
CC prodn. of an enzyme with improved properties.
SQ Sequence 2010 BP; 292 A; 756 C; 288 T;

Query Match 58.4%; Score 903; DB 1; Length 2010;
Best Local Similarity 81.0%; Pred. No. 4.7e-111;
Matches 1225; Conservative 0; Mismatches 265; Indels 23; Gaps 14;

Qy 2 acgagcgccttggtgactgggtgacagagtcacacacgacagcagcggaccctcgcc 61
Db 407 ACCAGGGCCTCTGGACTGGGTGACGTGTCCACTCCGACACGA-GCGGGCCCTCGC 465
Qy 62 tactacatcgctcctcctcttttccgctcagggcctcgacctgagcgtcagc 121
Db 466 TTCTGACAT-GGGTCTCTCCCTCTCCGCGGCTCAGCCCATGTGCTGNGGGCGCTCC 524
Qy 122 tatgcccggcctgtgggcccgggtgagacccggccgcccgttttctgttcctcggt 181
Db 525 GAGGCCCCGAGGCGTCCGTGGAGACATCTCTCGAAGACACCTTTGTGTCGAAGACTA 584
Qy 182 tcccttccagagagcgcgtcg--gcatactaatgttaatacgccttcagcgaataagt 238
Db 585 CCTTGCTCTTCCAGAGCGGTGGTGCACTACTTGTATAA-CGCCATGACGAATAAGT 643
Qy 239 cgaagcagcgaagcgcgcgtgactgaactaccagcccccagagagaggttcacc 298
Db 644 CGTACG-GAACAAGAGGAGCG--GCATGAGCTTCCAGCCCCACCCCGAGCAAGTTAC 701
```

```
Qy 299 ttcggcctgtgacccgtcggtgagggcgagcccttcggcgacgccaacgcgtccc 358
Db 702 TTCGGTCTGTGAGCGTGGCGTGGCAGGAAAGGACCCGTTTCGGCGACGCCACCGCCCT 761
Qy 359 gccctgacccgggtoga--cgtcagcggctggccgaactggcctacgaggtgacc 415
Db 762 GGCCTGACCCCGGTGAGACCGTGCAGCGCTGGCCGAGCTGGCGCCTACGGAGTGACC 821
Qy 416 ttcacagcagcagcctgatcccttcggggcgctccacacacgagcggagggcacgtc 475
Db 822 TTCACGACGACGACCTGAATCCCTTCGGGTCTGCCGACACCGACGCGAGTTCGCACATC 881
Qy 476 aagcggttccgtcagggcgtcgacgagcagcgcctgacccgttcogattggccaccaccac 535
Db 882 AAGCGTTCCGCCAGGCCCTGGACGCCACTGGCATGACGGTGGATGGCCACACGAAAC 941
Qy 536 ctcttaccaccaccccgcttctcaaggcagcgcgttcaacgccaacgacgcgagtgagc 595
Db 942 CTCTTACCCACCCCGTCTTCAAG---GACCGTTTCAACGCCCAACGACCGGAGCTGGCG 998
Qy 596 cgttacgcctcgcaagacacatccggaacatcgatctcgcggtcgagctggggcccaag 655
Db 999 GCCTAGCGCGTGGCAAGACCATCGGCAACATCGACCTGGCGGCGGAGCTGGGCGCCAG 1058
Qy 656 gtctacgtcgcctggggcgccgagggcgcgagtcctcgagtcggtgcgcgaagagcgtggt 715
Db 1059 ACGTACGTGCGCTGGGGCGCGGTGAGGGCGCGAGTCCGCTGGCGCAAGACGAGTGCGC 1118
Qy 716 gggcctcgacccgcaggaaggccttcgaacctgctcgcgagtagctacacctcgag 775
Db 1119 GACGCCCTCGACCGCATGAAGGAGGCGTTTCACCTCTTCGCGGAGTACGTACCCGCCCA 1178
Qy 776 ggctacgacatccggttcgcctcagagcccaagcgaacgagcgcgcgcgcgacatcgt 835
Db 1179 GGCTAGGACCTTCGCTTCGCATCGAGGCCCAAGCCCAAGACGAGCCCGGCGGAGATCCTC 1238
Qy 836 ctgccaccatcgccacgcgccttccttcagcgcctggagcgcctggagcgcctgagctgac 895
Db 1239 CTGCCACCGTGGCCACGCGCTTCACCTTCATCGAGCGCTGGAGCGCCGCGGAGTGTAC 1298
Qy 896 ggtgtaaccccgaggtgggccaagcagagcagatggccgacctgaacctcccgacagcgc 955
Db 1299 GGCGTCAACCCGAGGTGCGGCACGAGCAGATGGCCGCCCTGAACCTTCCCGACGCGCATC 1358
Qy 956 gcgcagcctctgtggcgggcaagctcttcacatcgacctcaacgagcgcgcgcgcgcgc 1015
Db 1359 GCGCAGGCCCTGTGGCGGGCAAGCTCTTCCACATCGACCTCAACGGCGAGTCCGGGCATC 1418
Qy 1016 aagtacgacccagagcctgcttcggcgccggtgacctgcgccgaccttctggtggtgctc 1075
Db 1419 AAGTACGACAGGACTGCGG--TTCGCGCGCGGACCTTCGCGGCGCTTCTGG-TGCTC 1475
Qy 1076 gacctgtggagagcgcgcgtggagggtcgcgcacacttcgacctcaagcccccgagg 1135
Db 1476 GACCTCTTGGAGAGCGCGGTTCAGAGGCCCGCGGCACTTCGACTTCAAGCCGCGCGG 1535
Qy 1136 accagagacatcgacgcgcgtgtggcctccgcgcgcgggtgcatgcgaactacatgac 1195
Db 1536 ACCGAGGACTTCGAGCGGTGTGGCCTTCGCGGAGGCTGCATGCGCAACTACCTGATC 1595
Qy 1196 ctgaaggagcgcgcgccttcctcgtcgaccccgagggtccagaggagc---cctggt 1252
Db 1596 CTCGAAGCAGC-CGCGCGCGCTTCGCGCGCACCGGAGGTGCAGGAGCGCGCTCGGC 1654
Qy 1253 gccgcgcgcctcgacacagctcgcgagacccaccgc-ggcgcagcgcctcgaggccctgct 1311
Db 1655 GCGCGCGCTCTGGACACAGCTGCGCCACCGCTGCGGACGCGCTGCGGAGGCGCTCTCT 1714
Qy 1312 ggcgcagcgcgcgcgtacgaggaacttcagctgagcgcgcgcgcgcgcgcgcgcgcgcgc 1370
Db 1715 CGCGACCGCAGCGGTTTCGAGGACTTCAGTGGAGCGCGCGCGCGCGCGCGCATGGT 1774
Qy 1371 tcgagcgcctcgacacagcctcgccatggaccacctgctggcgcccgcggtgacccgggc 1430
```



```

|||||
Db 1775 TCAGACGCTCGACAGCTGGCGATGGACACACTGTGGCGCGCGCGGTGACGTCGC 1834
QY 1431 gacgagggggtacgcggtgcgtatccctcgcctcatgaggggtgctggcggtc 1490
Db 1835 GCCAGGAGCGGTGACGGAAGACCGGGGACCGCGCGCTGCGCGGGCGTGTCTCC 1894
QY 1491 gaggcgccccgc 1503
Db 1895 CCGCGCGTCCGC 1907

RESULT 5
Q01862
ID Q01863 standard; DNA; 1567 BP.
AC Q01863;
DT 20-AUG-1990 (first entry)
DE Sequence of the rightward promoter (Pr) of bacteriophage lambda and
DE Actinoplanes missouriensis D-glucose isomerase (91) gene
KW D-xylose isomerase; D-xylose ketol-isomerase; plasmid pMas-GI;
KW plasmid pMc5-GI; transcriptional unit.
OS Bacteriophage lambda and Actinoplanes missouriensis.
EH Key Location/Qualifiers
FT cds 189..1373
FT /*tag=
FT misc_feature 137..1525
FT /*tag= b
FT /*note="DNA of Actinoplanes missouriensis"
FT EP355039-A.
PN 21-FEB-1990.
PD 17-JUL-1989; 201893.
PF 17-JUL-1989; EP-201893.
PI (KORN) Gist-Brocades NV (PLAN-).
PI Mrabet N, Lasters I, Stanssens P, Mathysens G, Wodak S, Quax WJ;
DR WPI: 90-052789/08.
PT Biologically active modified protein prodn.
PT comprises substituting arginine residue for lysine in starting protein
PS ; Fig 1b; 64pp; English.
CC Its starting protein is native D-glucose isomerase. It is a portion of
CC vectors pMas5-GI and pMc5-GI used in the process of making a modified
CC D-glucose isomerase gene. The method involves substituting one or more
CC Lys with Arg, or vice versa, at a site which can accommodate such a
CC substitution without altering the biological activity. The resulting
CC gene has enhanced thermostability and increased resistance towards
CC chemical modification.
SQ Sequence 1567 BP; 300 A; 532 C; 490 G; 245 T;

```

Query Match 38.8%; Score 599.6; DB 1; Length 1567;

Best Local Similarity 71.0%; Pred. No. 2e-71;

Matches 861; Conservative 0; Mismatches 324; Indels 27; Gaps 4;

```

QY 271 ccagccaccacccagagacaggttcaccttcgacctgacgctgagcggtggcg 330
Db 197 CCAGGCCACACCGGACAGACAAAGTTCCTTCGGTCTCTGACCGTGTGGATGGCAGCTCG 256
QY 331 ggacccttcgagcagcagcagcgtcccgccctcagaccggtcgga---cgtgagcggt 387
Db 257 TGACGGCTTCGGTGACGCCACGCGTACGGCACTCGACCCGGTCGAGGCCGCTGCACAGCT 316
QY 388 ggcgaactggcgctacgagtgaccttcacagcagcagcagcagcagcagcagcagc 447
Db 317 CGCTGAGATCGGGCGCTACGGCATCAGTTCACGATCCACGACGACGACCTGTGCTCCCTC 376
QY 448 gtccgacccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 507
Db 377 GGACGCCACAGACCGCGACGGCATCATCGCGGGCTTCAAGAGGCGCTCGACGAGACCGG 436
QY 508 catgaccttcgagtgccacacacacacacacacacacacacacacacacacacacac 567
Db 437 CTTGATGCTCCCGATGGTGACACCAACCTCTTCACCCACCCCGGTGTTCAAGGACGGCG 496

```

```

QY 568 gttcaacgccaacgacccgagcgagtgctgacgcttaccgcttgcgaagacacatccggaacat 627
Db 497 CTTACACGACACGACGCTTCGCTGGCGGTACGCGATCCGCAAGGTGCTGCGCCAGAT 556
QY 628 cgaatcgcggtcgagctggcgcaaggtctacgctcgctggggcgccgagagggcg 687
Db 557 GGACCTCGGGCGGAGCTGGGCGCGAAGAGCTGCTCTCTGGGGCGCGCGAGGGCGC 616
QY 688 ggaagtcggtgcccgaaggaagcgtgctgcccctggaccatgaaggaagccttcga 747
Db 617 CGAGTACGACTCGGCCAAGGAGCTAGCGCGCCCTCGACCGCTACCGCGAGCGCTCAA 676
QY 748 cctgctcgcgagtaagtcacctcgaggggtacgacatccggttcgcatcgagccaa 807
Db 677 CCTGCTCGCGCAGTACTCCGAGGACCGCGTTCAGGCGCTCGCTTCGCATCGAGCGAA 736
QY 808 gccgaacgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 867
Db 737 GCGGAACGAGCCCGCGCGGAGATCTGTCGACCGCGCGGCGACGCTCGCTTCGT 796
QY 868 cgaagcctggagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 927
Db 797 GCAGGAGCTGGAGCGTCCGAGCTCTTCGCGCATCAACCGGAGACCGGCGCACGAGCAGAT 856
QY 928 ggcggcctgaacttcccgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 987
Db 857 GTCGAACCCTCAACTTCACCCAGGGCATCGCCAGGCGCTGTGGCACAAAGAGCTGTCCA 916
QY 988 catcgacctcaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1047
Db 917 CATCGACCTGAACGGTTCAGACGCGCGCGGAGTTTCGACGAGGACCTGCTTCGCGCCACGG 976
QY 1048 tgacctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1092
Db 977 TGACCTGCTCAACGCGTTCGCTGCTGACCTCTCTGGAGAACGCGCGCGACGCGCGCC 1036
QY 1093 cggctgggaggtccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1152
Db 1037 GCGGTACGACGCGACCGCGTTCCTTCGACATACAAAGCGCTCCCGTACCGAGGACTACGACGG 1096
QY 1153 cgtgtgggctccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1212
Db 1097 CGTCTGGAGTGGCGAAGCCCAACATCCGATGTACTCTGCTCAAGGAGCGGGCCAA 1156
QY 1213 cgccttcctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1272
Db 1157 GCGCTTCGCGCGCGACCGCGAGGTGCGAGGCGCTCGCGCGCAGCAAGGTTCGCGAGCT 1216
QY 1273 -----cgccgagccacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1326
Db 1217 GAAGACCGCGACCTGAACCCGGCGGAGGATACGCCGAGCTGCTCGCGACGCGAGCGC 1276
QY 1327 gtacgaggaacttcgacgtggagcgc---ggccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1383
Db 1277 GTTCGAGGACTACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1336
QY 1384 ccagctcgccatggacacacacacacacacacacacacacacacacacacacacacacac 1443
Db 1337 CCAGCTCGCGATCGAGCACTGCTCGGAGCGCGCGCTGACCATGCGCTTGGTAGCGCGGATC 1396
QY 1444 ggcggtcgatc 1455
Db 1397 GACAGCTCGACC 1408

RESULT 6
Q01862
ID Q01862 standard; DNA; 1185 BP.
AC Q01862;
DT 20-AUG-1990 (first entry)
DE Nucleotide sequence of D-glucose isomerase (EcoAmi(DSM) GI) from
DE Actinoplanes missouriensis (DSN 43046) produced by engineered E. coli
KW Actinoplanes missouriensis (DSM 43046); D-glucose isomerase;

```



Query Match 38.7%; Score 598.4; DB 1: Length 1182;  
Best Local Similarity 72.1%; Pred. No. 3e-71;  
Matches 846; Conservative 0; Mismatches 301; Indels 27; Gaps 4;

Qy 271 ccagcccccacccgagcagggttcacccttgccttgtagcccgctggcgagggcg 330  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
9 CCAGGCCACACGCGAGACAGATTCTCCTTCGTCTCTGGACTTGGATGGCAGGCTCG 68  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 331 ggacctttcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 387  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 69 TGACGGCTTCGGTGAGCCACGCGTACGGCACTACGCACCTCGACCCTCGAGCGCGTGACAAGCT 128  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 388 ggcgaactggcgctacgaagtacgttacccttcaagcagcagcagcagcagcagc 447  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 129 CGGTGATCGGCGCTTACGGCATCACGTTCCACGACGACGACCTGGTGCCTTCGGCTC 188  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 448 gtccgacacccgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 507  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 189 GGAGCCCCACCCCGCAGCGCATCATCGGGCTTCAGAAAGGCGCTCGACGAGACCGG 248  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 508 catgaccttcgatggccaccaccaactcttcaaccaccocogtcttaagcagcagc 567  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 249 CTTGATCGTCCGATGTCACCAACCACTTTCACCCACCCCGGTTCAGAGACGCGG 308  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 568 gtccagcccaacgacccgacgacgacgacgacgacgacgacgacgacgacgac 627  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 309 CTTACACGACGACGCTTCCGTGGCGCTACGGCATCCGGAAGGCTGCGCCAGAT 368  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 628 cgatctcggctgagcgtggcgccaaggtctacgtcgcctggcgcgcgagggcg 687  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 369 GGACCTCGGCGCGAGCTGGCGCGAAGAGCGCTCGTCTCTGGGCGCGCGAGGCGC 428  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 688 ggagtcggtgccccaaagacgtcgctggcgccctggaccgcagcagcagcagc 747  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 429 CGAGTAGACGCTCGGCAAGACGCTACGCGCGCCCTTCGACCGCTACCGGAGCGCTCAA 488  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 748 cctgctcggagtagctcaectcgcagggctacgacatccggttcgcatcgagcccaa 807  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 489 CTGTCTCGCGAGTACTCCGAGGACCGGGTTACGGCTTCGGCTTCGCCATCGAGCCGAA 548  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 808 gccgaacgagccgccccgacatcctgtgccaccatcggccacgctgccttcac 867  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 549 GCCGACGAGCCCGCGGCGACATCTGCTCCGACCGCGCGCCACGCGCATCGCGTTCGT 608  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 868 cgagcctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 927  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 609 GCAGGAGCTGGAGCGTCCCAGGCTCTTCGGCATCAACCCGAGACCGGCGACGAGCAT 668  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 928 ggcggcctgaattccgcagcagcagcagcagcagcagcagcagcagcagcagc 987  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 669 GTGAACCTCACTTACCCAGGGCATCGCCAGGCGCTGTGGCACAAGAAGCTGTCCA 728  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 988 catgacctcaagcgcagtcggcgcacatgaagcacccagcagcagcagcagcagc 1047  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 729 CATGACCTGAACGGTCAGCACGCGCGGAGGTTTCGACACGAGGCTGTCTTCGCGCACGG 788  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 1048 tgacctgcgcgccgcttctggtgctgacctgtgtagagcg-----c 1092  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 789 TGACCTGCTCAACCGGCTTCGCTGTGTCAGACCTTCCTGGAGAACGCGCGCGCGCC 848  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 1093 cggtggaggggtcccgccacttcactcaagcccccgcgagcagcagcagcagcagc 1152  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 849 GGGGTACGACGACCCCGTCTCAGCTACAGACCGCTCCGTTACCGAGGACTACGACGG 908  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 1153 cgtgtggcctccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1212  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 909 CGTCTGGGAGTCGGCAAGCCCAACATCCGGATGTAACCTGCTGCTCAAGGAGCGGCCAA 968  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 1213 cgcttcctgctgcagcccgaggtccagagggccctcgctgcccgcgcgcgcgcgcgc 1272  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 969 GGGGTTCGCGGACCCCGAGGTGAGAGGGCGCTCGCGCGCGAGCAAGGTTCGCGAGCT 1028  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||











PA (HOFF) HOFFMANN LA ROCHE & CO AG F.  
PI Reeves CD, Soliday CL;  
DR WPI: 97-538615/50.  
DR P-PSDB: W34193-W34219.  
PT Streptomyces frenolicin gene cluster - useful for producing recombinant frenolicin antibiotics  
PS Claim 1: Page 40-60; 66pp; English.  
CC This DNA sequence comprises the Streptomyces frenolicin gene cluster containing specifically claimed coding sequences (genes A-U) that respectively encode 21 proteins (see W34199-219) involved in frenolicin synthesis. The genes can be divided into 5 subclusters: (1) genes A, B, C, D and F encode an efflux pump; (2) genes H, I, J and K encode butyrate starter synthases; (3) genes L, M and N encode polyketide synthases (PKS); (4) genes O, P, Q and R encode a hemiketalase, a ketoreductase and cyclases/dehydrases; CC and (5) genes S and T encode a keto/enoyl reductase and a hydrolase. Also claimed are vectors, host cells (especially a Streptomyces sp., particularly Streptomyces roseofulvus), and the encoded proteins. Cells transformed using the above sequence can be cultured to produce frenolicins or frenolicin precursors. The precursors can be converted to frenolicins by chemical or other methods. The frenolicins can be oxidised to frenolicin B, an antibiotic used as an anticoccidial agent. The frenolicins can be used as animal feed additives.  
SQ Sequence 24379 BP; 3077 A; 9792 C; 8499 G; 3011 T;

Query Match 6.0%; Score 93; DB 1; Length 24379;  
Best Local Similarity 44.5%; Pred. No. 1.5e-05;  
Matches 415; Conservative 0; Mismatches 515; Indels 3; Gaps 1;

QY 484 ccgtcagcgctcgcagcgacggcgatgaccgttcgatggccaccacacacattcttcac 543  
DB 15033 CCGGTCGGCTCGTCGCCGCCGACGAGGACACTGAAGGACGGCGGCTTCGACCCGGCCCG 15092  
QY 544 ccaccccgctttcaaggaagggcggttcacccgacgacggcgagtcgctgcgttaagc 603  
DB 15093 GCTGCCGAGTACGGGCGGAGCGCGTCACTCCACAGCGCACCGCGGCTTCGAGTTTAC 15152  
QY 604 cctggcaagacacatccggaac---atcgatctcggctcgagtcggcgcaaggtcta 660  
DB 15153 CCACGCGGAGATCCCGAAACTGTGACCGAGGGCGCGCGCGTCCAGCGTCTACGAGTC 15212  
QY 661 cgtcgcctggggcgcgcgagggcgagtcgctgctgcgcgaagagtcgctgcgcgc 720  
DB 15213 CTTGCGCTGGTTCTACGCGCGTCAACACCGCGCAGATCTCCATCCGCGACGGATCGCGGG 15272  
QY 721 cctgagacgcataagggagggccttcgacctgctggcgagtcgtaacctcgacgggcta 780  
DB 15273 CCGCGCGCGCTCGTGTGGCGGACGAGCGGGCGGCTTCGACGCGCTCGGCGACGGCCG 15332  
QY 781 cgaacatcggttcgcacatagaccccaagcgaacgagcgcgcgagacatctactgccc 840  
DB 15333 CCGCGTACTGCGCAAGGCGGGGTGCTGGCGGTGAGCGGCGGTGGAGTCCGCGCTCGA 15392  
QY 841 caccatcgcgacgcgctgccttcacgcagcgcctggagcgcccccgagctgacgggtg 900  
DB 15393 CCGCTGGGCGCTTGGCGGCCCGCCCGCTGTCGGGACCCCTCAGCGGCTCGCGGACCGCGC 15452  
QY 901 caacccccaggttggccacagcagatggccggttcgtaacctccgcacgagcatcgca 960  
DB 15453 CACCGCGTACCTCCGCTTGCAGCGGCGTCCCTCGGACCGCTCGTGGGAGGGGGCGCGC 15512  
QY 961 ggcctgtgtggcggaagccttccacatgcacatcactcaacgcccagtcgagtcacagta 1020  
DB 15513 CTTCTTCACTCCCTGGAGACACCCCGGACCGCCGAGGACCGACCGACCCCGGATCTACGG 15572  
QY 1021 cgaccagggacctgcttggcgcgcggtgacctgcgcgcgcttccttggctgtgacact 1080  
DB 15573 CGAACTCGCGGGTAGCGCGCGCACGTTCCACCCCGCGGGCTCCGAGCGGCCCGCCCG 15632  
QY 1081 gctgagagcgccgctgtggaggggtccgcgcaacttcgacttaagcccccgcggagcaga 1140

DB 15633 CCTCGAAGCGCGCGCGCTCGCTCGCGACGCGGGCTGGCACCCCGGGGAGCTCGA 15692  
QY 1141 ggacatcgacggcggtgtggcctcgcgcgcggtgcatggaactacctgatacctgaa 1200  
DB 15693 CGTCTCTTCGGGACGCGGGGGGCTCCCGCGCCCGCCGACGCGCGCGCGCCCT 15752  
QY 1201 ggagcgcgcgcgcttccgtgcccgaaccccgaggtccagggagggccctgctgctgcgcgcg 1260  
DB 15753 GCGCGCGCTCTCGGCCCGCGGGCGGTTCGGGTGAGCTGCCGAAGACCCAGACCGGGCG 15812  
QY 1261 gctgacacagctcgccgagccacccgcgccgagccctgcaaggccctgctgctgcgcgcg 1320  
DB 15813 GCTCGCTCGGGCGCGCGCGCTCGAGCTGCGCCCGCCCTCTCTCGCGCTCGCGGAGGG 15872  
QY 1321 caccgcgtacggaggttcgacgtgagcgcgcgccgcgcgccatgcttcgagcgctc 1380  
DB 15873 CTTCTGCCCCCGCGCTCCACCTCGAGAGTCCGCGGTACGGGCTCGAGCTCGT 15932  
QY 1381 cgaccagctcccatggaccacctgctggcgcg 1413  
DB 15933 CGCGGACACCGCGCGCGCTCCCGCTGGGCAC 15965

RESULT 13  
ID V25925 standard; cDNA; 24379 BP.  
AC V25925;  
DT 15-JUL-1998 (first entry)  
DE Streptomyces roseofulvus frenolicin gene cluster.  
KW Streptomyces roseofulvus; frenolicin gene cluster; frenolicin B; antibiotic; ss.  
OS Streptomyces roseofulvus.  
FH Key Location/Qualifiers  
FT CDS 636..2948  
FT /tag= a  
FT /transl\_except= (pos:636..638,aa:Met)  
FT /note= "encodes protein given in W55800"  
FT 2945..3916  
FT /tag= b  
FT /transl\_except= (pos:2945..2947,aa:Met)  
FT /note= "encodes protein given in W55801"  
FT 4020..4844  
FT /tag= c  
FT /transl\_except= (pos:4020..4022,aa:Met)  
FT /note= "encodes protein given in W55802"  
FT 4841..6415  
FT /tag= d  
FT /note= "encodes protein given in W55803"  
FT 6533..7183  
FT /tag= e  
FT /transl\_except= (pos:6533..6535,aa:Met)  
FT /note= "encodes protein given in W55804"  
FT 7344..8897  
FT /tag= f  
FT /note= "encodes protein given in W55805"  
FT 9164..10012  
FT /tag= g  
FT /transl\_except= (pos:9164..9166,aa:Met)  
FT /note= "encodes protein given in W55806"  
FT complement (10105..10621)  
FT /tag= h  
FT /transl\_except= (pos:10621..10619,aa:Met)  
FT /note= "encodes protein given in W55807"  
FT complement (10618..11628)  
FT /tag= i  
FT /transl\_except= (pos:11628..11626,aa:Met)  
FT /note= "encodes protein given in W55808"  
FT 11809..12066  
FT /tag= j  
FT /note= "encodes protein given in W55809"  
FT complement (12154..13209)  
FT /tag= k  
FT /transl\_except= (pos:13209..13207,aa:Met)

[illegible]





THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2000, 23:05:43 ; Search time 44.32 Seconds  
(without alignments)  
4176.263 Million cell updates/sec

Title: US-09-383-318-1  
Perfect score: 1 cagcagccttggtgact.....gcgcggtgtgggcgcgtgc 1546  
Sequence:

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 214294 seqs, 59861574 residues

Total number of hits satisfying chosen parameters: 428588

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/5C\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/5D\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/6\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PCUS9\_COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	952.8	61.6	1164	1	US-07-640-476-6
2	599.6	38.8	1566	7	Sequence 6, Appli
3	598.8	38.7	1185	7	Patent No. 5290690-3
4	598.4	38.7	1182	1	Patent No. 5290690-1
5	481.8	31.2	1805	1	Sequence 4, Appli
6	98.6	6.4	1820	1	Sequence 1, Appli
7	97	6.3	1821	3	Sequence 7, Appli
8	85.2	5.5	44377	3	Sequence 7, Appli
9	85.2	5.5	44377	3	Sequence 7, Appli
10	82.8	5.4	4257	3	Sequence 1, Appli
11	79	5.1	1668	4	Sequence 1, Appli
12	78.6	5.1	1908	1	Sequence 1, Appli
13	78.6	5.1	1908	3	Sequence 1, Appli
14	78.6	5.1	12001	2	Sequence 11, Appli
15	77.4	5.0	30001	1	Sequence 1, Appli
16	77.4	5.0	30001	3	Sequence 10, Appli
17	77	5.0	1620	3	Sequence 10, Appli
18	76.4	4.9	1998	7	Patent No. 5212296
19	76.4	4.9	5392	3	Sequence 1, Appli
20	75.4	4.9	833	3	Sequence 3, Appli
21	75.2	4.9	12588	4	Sequence 1, Appli
22	73.6	4.8	2522	1	Sequence 1, Appli
23	72.2	4.7	2668	3	Sequence 11, Appli
24	72	4.7	1931	4	Sequence 2, Appli
25	72	4.7	20235	2	Sequence 3, Appli
26	71.2	4.6	8438	1	Sequence 1, Appli
27	71	4.6	1146	2	Sequence 1, Appli

28	71	4.6	2728	2	US-08-482-385A-5	Sequence 5, Appli
29	70.4	4.6	1268	3	US-08-403-852D-2	Sequence 2, Appli
30	70.4	4.6	2312	1	US-07-736-178C-1	Sequence 1, Appli
31	70.4	4.6	43280	3	US-08-804-227C-1	Sequence 1, Appli
32	70	4.5	2721	7	5215881-2	Patent No. 5215881
33	69.2	4.5	1208	3	US-08-403-852D-4	Sequence 4, Appli
34	68.8	4.5	1771	3	US-08-533-669A-7	Sequence 7, Appli
35	68.8	4.5	1771	4	US-08-511-872-1	Sequence 1, Appli
36	68.8	4.5	11219	2	US-07-642-734C-1	Sequence 1, Appli
37	68.6	4.4	3468	3	US-07-951-715A-2	Sequence 2, Appli
38	68.6	4.4	3468	1	US-08-459-449A-2	Sequence 2, Appli
39	67.2	4.3	4257	3	US-08-690-473-1	Sequence 1, Appli
40	66.2	4.3	1288	2	US-08-440-856A-9	Sequence 9, Appli
41	66	4.3	2864	1	US-08-343-428-1	Sequence 1, Appli
42	65.6	4.2	30001	1	US-08-125-468-1	Sequence 1, Appli
43	65.6	4.2	30001	3	US-08-474-933-1	Sequence 1, Appli
44	64.4	4.2	4776	4	US-08-852-401-1	Sequence 1, Appli
45	63.8	4.1	8051	4	US-08-576-626A-2	Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-07-640-476-6  
; Sequence 6, Application US/07640476  
; Patent No. 5376536  
; GENERAL INFORMATION:  
; APPLICANT: QUAX, WILHELMUS  
; APPLICANT: LUITEN, RUDOLF G.M.  
; APPLICANT: SCHUURHUIZEN, PAUL W.  
; APPLICANT: MRABET, NADIR  
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASE ENZYMES AND  
; TITLE OF INVENTION: THEIR USE  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 545 Middlefield Road, Suite 200  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/640,476  
; FILING DATE: 19910110  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kate H. Murashige  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 24615-20009.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 327-7250  
; TELEFAX: (415) 327-2951  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1164 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptomyces murinus  
; STRAIN: DSM 40091  
; FEATURE:

```
; NAME/KEY: CDS
; LOCATION: 1..1164
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /EC_number= 5.3.1.5
; OTHER INFORMATION: /product= "xylose isomerase (glucose isomerase)"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /standard_name= "D-xylose ketol isomerase"
US-07-640-476-6

Query Match      61.6%; Score 952.8; DB 1; Length 1164;
Best Local Similarity 89.9%; Pred. No. 8.8e-152;
Matches 1046; Conservative 0; Mismatches 112; Indels 6; Gaps 2;

QY 263 atgaactaccagccacccccagagagaggttcaccttcggcctgtgacccgtgctgg 322
DB 1 atgacgttcacagccacccccagagagaggttcaccttcggcctgtgacccgtgctgg 60
QY 323 cagggcgggaccccttcggcgagccacgcgctgcgcgccttcgaccccggtcga---cg 379
DB 61 CAGGGAAGGGACCCGTTTCGGCGACGCCACCGCCGCCCTCGACCCGGTCGAGACGGTG 120
QY 380 cagcggtcgccgaactggcgccctaacgagtgagtgaccttcacagacgacctgacccc 439
DB 121 CAGCCGCTGGCGAGCTGGCGCCTACGAGTGACCTTCCACGACGACGACCTGATCCCC 180
QY 440 ttcggggcgctcgacacccagcgagcgagcgcaecgctcaagcggttcctcagcgctgcac 499
DB 181 TTCGGGTCTCTCGGACACCGAGCGCGAGTGCACATCAAGCGGTTCGGCAGGCCCTGGAC 240
QY 500 gcagccgcatgacggttcctcgatggccaccacccaccttcacccaccccgcttcgaag 559
DB 241 GCCACCGCATACGCTGCCCCATGCCCCACCAACCACTTTCACCCACCCGCTTCAAG 300
QY 560 gcagcgcggttcacgcgcacacgcgcgcgcagtcgcccgttaagccctgcgcgaagaccatc 619
DB 301 GACGGCGGCTTACCGCGCAACGACCGCGACGTACGCCGGTACGCCGTGCGCAAGACGATC 360
QY 620 cggacacatcatctcgcggtcgagctggcgccaaaggttcctagctgcctggggcgccgc 679
DB 361 GGCACATCGACCTGGCGGCGCAACTGGGTGCCAGAGCATGTGCGCTTGGGCGGCGGT 420
QY 680 gagggcgcgagtcggtgcccgaagacgtgctgvggcctggaacgcgcatgaagag 739
DB 421 GAGGGCGCGAGTTCGGGTGGCCCAAGGACGTGCGGACGCCCTCGACCGCATGAAGAG 480
QY 740 gcttcgacctgctggcgagtagctacacctgcagggcgtaagacatccggttcgcatc 799
DB 481 GCGTTTCGACCTCTCGCGGAGTACGTACCGGCCCAAGGGCTACGACCTCCGCTTCGCCATC 540
QY 800 gagcccaagcgaaagcgcgcgcgagacatcctgctgcccacacatcgccacgcgtc 859
DB 541 GAGCCCAAGCCCAAGACGCCCGCGGCGACATCTCTGTCGCCACCGTGGCCACGCCCTG 600
QY 860 gcttcacagcgcctggagccccagctgtgtacggtgtcaaccccagagtgggccac 919
DB 601 GCTTCATCGAGCGCTGGAGGCCCGGAGCTGTACGGCGTCAACCGAGGTGCGGCCAC 660
QY 920 gagcagatggccgacctgaacttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 979
DB 661 GAGCAGATGGCGGCGCTGAACCTCCCGCACGCATCGCGCGAGCGCCCTGTGGCGGGCAAG 720
QY 980 ctctccacatcgacctcaacggccagtcggtcgatcaagtagcagcaggaacctcgcttc 1039
DB 721 CTCCTTCACATCGACCTCAACGGCCAGTCCGGCATCAAGTAGACACGACCTGCGGTTTC 780
QY 1040 ggcgcggtgacctcgcgccgcttcctggtggttcgacctgctggagagcgccgctgg 1099
DB 781 GCGCGCGGCGACCTGCGGGCGCGCTTCTGGCTGGTCTGACCTCTCTGAGACCCGCGTTAC 840
QY 1100 gagggctcgccgcaacttcgaccttaagcccccgcggaacccgaggaacatcgagcgctgg 1159
DB 841 GAGGCCCCCGCGCACTTCGACTTCAAGCGCGCGGACCGGAGGACTTCGACGCGGTGG 900
```

```
QY 1160 gcctccgcgcgcggtgcatcgcaactacatgatcctgaaggagcgcgccgccttc 1219
DB 901 GCTTCGGCGCGGCTGCATGCGCAACTACCTGATCTCAAGGACGCTGCGGCGCCTTC 960
QY 1220 cgtgccaccccgaggttcacagagagccctcgctgcgcgcgcgcgcgcgcgcgcgcgcgc 1279
DB 961 CGTCCGACCCGAGGTGACGAGGCGCTGCGTGGCGGCTCTGGACACGCTGGCCAG 1020
QY 1280 cccaccgcgcgagcgcgctcgagccctgctggccgacccacacgcgcgtacagagacttc 1339
DB 1021 CCGACCGCGCGGACGCGCTTTCAGCCCTGCTCGCCGACCGCGGCTTCGAGGACTTC 1080
QY 1340 gacgtgga---cgcgcgcgcgcgcgcatgaccttcgagcgccctcgacacgctcgccatg 1396
DB 1081 GACGTGATGCGCGCGCGCGCGGTATGGGTTCGAGACCTCGACCACTGGCGATG 1140
QY 1397 gaccacctgctggcgcccgcgccg 1420
DB 1141 GACCACCTGCTGGCGCGCGCGGC 1164

RESULT 2
5290690-3
; PATENT NO. 5290690
; APPLICANT: MRABET, NADIR; LASTERS, IGNACE; STANSSENS, PATRICK
; MATTHYSSENS, GASTON; WODAK, SHOSHANA; QUAX, WILHELMUS J.
; TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
; STABILITY OF PROTEINS
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/398.706
; FILING DATE: 25-AUG-1989
; SEQ ID NO.3:
; LENGTH: 1566
5290690-3
```

```
Query Match      38.8%; Score 599.6; DB 7; Length 1566;
Best Local Similarity 71.0%; Pred. No. 8.2e-93;
Matches 861; Conservative 0; Mismatches 324; Indels 27; Gaps 4;
```

```
QY 271 ccagccaccccccagagagaggttcaccttcggcctggtgacccgtcgctggcagggcg 330
DB 196 ccaggccacacgcgaagacaagtctctctgctctgacctgttgatggcagggctcg 255
QY 331 ggaccccttcggcgagccacgcgcgtcccgccctcgaccccgctga---cgcgagcggt 387
DB 256 tgacgcttcggtgacgcgcgcgcgtacggcactcgaccccggtcgagcgctgcaacgct 315
QY 388 ggccgaactggcgctacgcgcgcgcgttcacgttcacagacgacacgtgatcccttcggggc 447
DB 316 cgtcgagatcgccgctacgcgcgcgcgttcacagacgacacgtggtgccttcggctc 375
QY 448 gtccgacaccgagcgagggcgacgctcaagtggttcgctcagcgctcgacgcgacccg 507
DB 376 ggagcccgagcccgcgagcgacgcgcgcgttcacgagggcgctcgacgcgacccg 435
QY 508 catgaccttcgatggccacacacacacacacacacacacacacacacacacacacacacac 567
DB 436 cctgategctccgctggtgacccacacacacacacacacacacacacacacacacacacac 495
QY 568 gttacccgcaacacacacacacacacacacacacacacacacacacacacacacacacac 627
DB 496 cttacacgaacacacacacacacacacacacacacacacacacacacacacacacacac 555
QY 628 cgatctcgctgagctggggcgacaggttatagctgcctggggcgccgagggcgccg 687
DB 556 ggacctcgccgagctggggcgagagacgctgctgctctggggcgccgagggcgccg 615
QY 688 ggagtcggttcgccaagagagcgctgctgcccctgacgcgcgcgcgcgcgcgcgcgcgcgc 747
DB 616 cgagtacgactcgccgaagagcgctgaagcgccgctcgacccgctaccgagggcgctcaa 675
```







Db 1089 GTTCGAGGACTACGAGCCGACGACCCCTGGCGCCAAAGGCTTCGGCTTCGTCAAGCTGAA 1148  
 Qy 1384 ccagctcgcacagaccacactgctggtgcccgcgc 1417  
 Db 1149 CCAGCTCGGATCGAGCAGCACTGCTCGGAGCCGCG 1182

RESULT 5

US-08-112-630-1  
 ; Sequence 1, Application US/08112630  
 ; Patent No. 5411886  
 ; GENERAL INFORMATION:  
 ; APPLICANT: UDAKA, Shigezo  
 ; APPLICANT: SAKAGUCHI, Kenji  
 ; APPLICANT: YAMAGUCHI, Hideo  
 ; APPLICANT: DEKKER, Koen  
 ; TITLE OF INVENTION: XYLOSEISOMERASE GENE OF THERMUS  
 ; TITLE OF INVENTION: AQUATICUS, XYLOSEISOMERASE AND PROCESS FOR PREPARATION OF  
 ; TITLE OF INVENTION: FRUCTOSE  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; STREET: George Mason Bldg., Washington & Prince Sts.  
 ; CITY: Alexandria  
 ; STATE: Virginia  
 ; COUNTRY: United States  
 ; ZIP: 22313-1404  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/112,630  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/783,150  
 ; FILING DATE: 28-OCT-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Rea, Teresa S.  
 ; REGISTRATION NUMBER: 30,427  
 ; REFERENCE/DOCKET NUMBER: 024705-007  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 836-6620  
 ; TELEFAX: (703) 836-2021  
 ; TELEX: 440 580  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1805 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 288..1448  
 ; US-08-112-630-1

Query Match 31.2%; Score 481.8; DB 1; Length 1805;  
 Best Local Similarity 62.5%; Pred. No. 3.8e-73;  
 Matches 788; Conservative 0; Mismatches 467; Indels 6; Gaps 2;  
 Qy 269 taccagaccaccccccagagacaggttcaacttggcctgtgaccctgctggtgagggg 328  
 Db 291 TAGAGCCCAAAACCGAGCAGCACTTACCTTTGGCCTTTGGACTGTGGCAATGTGGGC 350  
 Qy 329 cgggaccccttcggcagaccacagctcccgccctcgaccctcgacatg---cagcgg 385  
 Db 351 CGTGATCCCTTCGGGACGCGGTTCGGGAGAGGCTGGACCCCGGTTTACGGTGTTCATAAG 410

Qy 386 ctggcgcaactggggcgcttacggagtgaccttccacgacgacgacgtgatcccccttcggg 445  
 Db 411 CTGGCGGAGACTTTGGGCGCTACGGGTAAACCTTACGACGAGGACCTGATCCGCGGGGC 470  
 Qy 446 gcttcgacacacgacgacgagcgacgacgttccgctcagcgctcagcgacgacgac 505  
 Db 471 ACCTCTCTCAGAGCGGAGCAGATCGTGTAGCGGCTTCAAGAAAGCTCTCATGAACCC 530  
 Qy 506 ggcagacccgttccgatggccaccacacacacacacacacacacacacacacacacac 565  
 Db 531 GGCTCAAGTCCCATGCTACCGGCCAACCTTCTTCGACCTGCTTTCAGGAGCGGG 590  
 Qy 566 gcttcacgcccacacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 625  
 Db 591 GCCTTACAGAGCCCGGAGCCCTTGGGTTCGGGCTATGCCCTTCGGAAGAGCCTTGGAGAC 650  
 Qy 626 atcgatctcgcgctgacgtggcgccaaaggtctacgtcgtcgtcgtcgtcgtcgtcgtc 685  
 Db 651 ATGGACCTGGGGCAGAGGTGGGGCCGAGATCTAGTGTCTGGCCGGCGGGGAGGGA 710  
 Qy 686 gcgaggtccggtgcccgaaggtgctgctgctgctgctgctgctgctgctgctgctgct 745  
 Db 711 GCTGAGGTGGAGGCCACAGGCCGCAAGGCCGGAAGGTCTGGGACTGGGTGCGGAGCGCTG 770  
 Qy 746 gacctgctggcgagtagctacacctcgagggctacgacacatccggttcgacctcagccc 805  
 Db 771 AACTTCATGCGCCCTACCGCAGGACAGGATACGGGTACCGGTTTCCCTTCGAGACC 830  
 Qy 806 aagcgaacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 865  
 Db 831 AGCCTAACGAGCCCGGGGGGACATTTACTTCGCCACCGCTGGGGGAGAGATCTCGCTTT 890  
 Qy 866 atcgagcgtcgtgagcgccccgagctgtacgtgtgtaacccccgaggtggcgacgagcag 925  
 Db 891 ATTATACCTGGACGCGCGGAGGCTTTCGGCTGAACCCGAGTTCGCCACGAGACC 950  
 Qy 926 atggcgccgctgaacttcccgcacgacgacgacgacgacgacgacgacgacgacgacgac 985  
 Db 951 ATGGCGGGCTTAACTTTGTCACGCGGTGGCCAGGCTCTCGACGCGGGAGAGCTTTTC 1010  
 Qy 986 cacatcgacctcaacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 1045  
 Db 1011 CACTTGACTCAACGACCAACGAGGTAGCGGCTTACAGAGGACCTCGCTTCGGCTCG 1070  
 Qy 1046 ggtgacctgcgcgcccttctggtgctgacctgctgacctgctgacctgctgacctgctg 1105  
 Db 1071 GAGAACTCAAGGGCGCTTTTCTGTGTGACCTCTCTGAAAGCTCCGCTACGAGGC 1130  
 Qy 1106 ccgcgccacttcgacttcaagcccccgacgacgacgacgacgacgacgacgacgacgacgac 1165  
 Db 1131 CCGCGCACTTTGACGCCACGCGCTGCTACCGAGGACGAGAGGGGTTTGGGCTTC 1190  
 Qy 1166 gcgcgccggtgcatcgcaactacgtacctgaagagcgacgacgacgacgacgacgacgac 1225  
 Db 1191 GCCGAGGCTGATGCTGCTACCTACCTGATCTTAAAGAAAGGCTGAAGCTTCCCGGAG 1250  
 Qy 1226 gacccgaggttcagagggccctgctgctgctgctgctgctgctgctgctgctgctgct 1285  
 Db 1251 GATCCCGAGGTCAAGGAGCTTCTTGGCGCTTACTATCAAGAGATCTTCGGGCTTGGCC 1310  
 Qy 1286 gcgcgacgagcgtcagggccctgctgctgctgctgctgctgctgctgctgctgctgct 1345  
 Db 1311 CTTTGGGCGCTTACTTCCCGGAGAGGCGGAGCCCTCAAGCGGGGAGGCTTCCCTC 1370  
 Qy 1346 gacgcgccgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 1402  
 Db 1371 GAGGCCAAAGCGCGCGGGGTTATGCCCTGGAACGCTGGACAGCTGGCGGTGGAGTAC 1430  
 Qy 1403 ctgctggtggtcgtgacccggtgacgacgacgacgacgacgacgacgacgacgacgacgac 1462  
 Db 1431 CTCCTGGGGGTGCGGGGTGAGGGGCGCCATCGGCTTGGACCTGGGAAGAGCGGGCTCA 1490  
 Qy 1463 gtcgtcatgaggggtgtgtggcggtcgtgagcgccccgccccccatcgtgtcgtctcc 1522



[illegible]

## RESULT

```

US-08-265-310-7
; Sequence 7, Application US/08265310
; Patent No. 5856166
; GENERAL INFORMATION:
; APPLICANT: Bartfeld, Daniel
; APPLICANT: Butler, Michael J.
; APPLICANT: Hadary, Dany
; APPLICANT: Jenish, David
; APPLICANT: Krieger, Timothy
; APPLICANT: Malek, Lawrence T.
; APPLICANT: Soostmeyer, Gisela
; APPLICANT: Walczyk, Eva
; APPLICANT: Krygsman, Phyllis
; APPLICANT: Garven, Shelia
; TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
; TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/265,310
; FILING DATE: 24-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,508
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 18740/133/CACO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 672 5300
; TELEFAX: 202 672 5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1821 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 104..1720
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 104..244
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 245..1720
;
US-08-265-310-7

```

Db 1536 CCCGACCGGACACCCCTACCGCTGGCGAGGCCCTCTCCGACAGCTCACCTCCG 1595  
 QY 1320 GCACCGGTGAGGAGCTGAGCTGAGCGCGCGCGCGGCGATGCTGAGCGCC 1379  
 Db 1596 GCCACCTCTCCTACCTAGAGGAGAGCGGCACACCGGTAGCGCGCGGAGCTCTTGA 1655  
 QY 1380 TCGACGAGTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1420  
 Db 1656 TCGACTCCGGGATCAACACGATACCTGCTGACCGGACCGCC 1696

## RESULT 8

US-08-804-227C-7  
 ; Sequence 7, Application US/08804227C  
 ; Patent No. 5876991  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dehoff, Bradley S.  
 ; APPLICANT: Kuntz, Stuart A.  
 ; APPLICANT: Rostock, Paul R., Jr.  
 ; APPLICANT: Sutton, Kimberly L.  
 ; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES  
 ; NUMBER OF SEQUENCES: 15  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: THOMAS G. PLANT 1501  
 ; STREET: LILLY CORPORATE CENTER  
 ; CITY: INDIANAPOLIS  
 ; STATE: IN  
 ; COUNTRY: USA  
 ; ZIP: 46285

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: MS-DOS  
 SOFTWARE: ASCII(DOS) Text only  
 CURRENT APPLICATION DATA:  
 FILING DATE: February 21, 1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Plant, Thomas, G.  
 REGISTRATION NUMBER: 35,784  
 REFERENCE/DOCKET NUMBER: X-8231  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 317-276-2459  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4437 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:

NAME/KEY: CDS  
 LOCATION: 350..14002  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 14046..20036  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 20110..31284  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 31329..36071  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 36155..41830  
 US-08-804-227C-7

Query Match 5.58; Score 85.2; DB 3; Length 4437;  
 Best Local Similarity 44.66; Pred. No. 6.7e-07;  
 Matches 333; Conservative 0; Mismatches 413; Indels 0; Gaps 0;

QY 671 ggccgcgcgagggcgcgaggtccggtgcccgaagagacgtgctgcccctggagccgc 730  
 Db 30004 GCCCGCGCGGGCGCGGACGCGAGGCGCGCGAGGTGTGCGCGACCTCGCGGCC 30063  
 QY 731 atgaagagagccttgagcctgctgagagtagctacgtcacctcagaggtctacgacatccgg 790  
 Db 30064 CTGGGCGGTGCGGCCACCGTGGTGGCTGGAGCTCTCCAGCGCGAGCGCTCCGCGGA 30123  
 QY 791 ttgcgcatcgagcccaagccgaagagcgcgcgagacatcctgctgcccacccatcgcc 850  
 Db 30124 CTGCTCGCGGCATACCGGCCGATCACCCCGCTGACGGCGGTGTGCACACACCGCGGTC 30183  
 QY 851 cagcgctgccttcctcatcgagcgccctggagcgccccagctgtactacgtgtcaaacccccgag 910  
 Db 30184 CTCGACGACGGCTGCTGCCCGGCTCACCCCGAGCGGATGCGGCGCTGCTGCGGCC 30243  
 QY 911 gtgggcaagcagagatggcgccctgaacttcccgcagcgccatcgcgaggtctctgtgg 970  
 Db 30244 AAGGTGGAGCGCGCGCTCCACCTGGAGAACTCACCGCGACCTCGACCTGTGCGCGTTC 30303  
 QY 971 gcgggcaagctcttccacatcgacctcaacgagccagtcggcgatcaagtagtaccgagagac 1030  
 Db 30304 GTCTCTTCTCTCCAGCGCGGTCTGCTGGGAGCGCCGCCAGGCAACTGACGGCGG 30363  
 QY 1031 ctgcgttcggcgccgtgagctgcgcgcccttctgtgctgagacctgctggagagc 1090  
 Db 30364 GCCAACGCCACCTCGACGCGCTCGCGCGCGCGCGGCTCCGCGCTCCGCTCGGTG 30423  
 QY 1091 gcgggctggagggctcgccacttcgaacttcaagcccccgagcgagggagacatcgac 1150  
 Db 30424 TCACCTCGCTGGGTGCTGTGTCCGACACAGCGGATGGCACACGACCTGACACGAG 30483  
 QY 1151 gggtgtggtcccgcgcgcggtgcatgcgaactacctgatcctgagagggcgcgcc 1210  
 Db 30484 AGCTCAGCGGGCTTCGCGCGGAGCGGCTTCGCGCGCTGTCGCGCACGCTGGGCGCC 30543  
 QY 1211 gccgcttcctgctccgagcccgaggtccagagggcctgctgcccgcgctcgagaccag 1270  
 Db 30544 GCGCTGTTCGACGCGCGCTCGGGTGGAGGAGCGGTGCGAGGTCCCATGCGGTTTCGAC 30603  
 QY 1271 ctgcgagagccacccggcgcgagcggcctgagggcctgctgcccagaccacccggtac 1330  
 Db 30604 CCGCGCGCGCTGCGCGCACCGGAAAGCGTCCCGCGCTGCTGTCGGACCTCGTCGGGTCC 30663  
 QY 1331 gaggacttcagctgagcgcgcgcgcgcgcgcatggccttcgagcgccctcgagcagctc 1390  
 Db 30664 GCCCGCGGACCGGTTCGCGCGCGCGCGCGCGCTTCGCGCGCTTCGCGGACCGCGG 30723

## RESULT 9

US-08-804-198-1  
 ; Sequence 1, Application US/08804198  
 ; Patent No. 5945320  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Burgett, Stanley G.  
 ; APPLICANT: Kuntz, Stuart A.  
 ; APPLICANT: Rao, Nagaraja R.  
 ; APPLICANT: Richardson, Mark A.  
 ; APPLICANT: Rostock, Paul R., Jr.  
 ; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: PAUL R. CANTRELL 1138  
 ; STREET: LILLY CORPORATE CENTER  
 ; CITY: INDIANAPOLIS  
 ; STATE: IN  
 ; COUNTRY: USA  
 ; ZIP: 46285

















THIS PAGE BLANK (USPTO)



Result No.	Score	Query Match	Length	DB ID	Description
1	82	5.3	925	79	CNS0091P
2	78.8	5.1	935	79	CNS006XK













[illegible]







Db 394 GTGKNTTGCVGSTCTTKSSCTBSSGTSBSTS 361

Search completed: March 22, 2000, 23:16:27  
Job time: 724 sec

THIS PAGE BLANK (85-70)

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: March 22, 2000, 23:30:54 ; Search time 21.4 Seconds  
(without alignments)  
239.853 Million cell updates/sec  
Title: US-09-383-318-2  
Perfect score: 2029  
Sequence: 1 MNYQPTDEFTGLTWGVG.....GMAFERLDQLANDHLLGARG 386  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 133990 seqs, 13297546 residues  
Total number of hits satisfying chosen parameters: 133990  
Minimum DB seq length: 0  
Maximum DB seq length: 1000000  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database : Issued Patents, AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/PCUS9\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	1900	93.6	388	1 US-07-637-870-4 Sequence 4, Appli
2	1900	93.6	388	1 US-07-637-399-7 Sequence 7, Appli
3	1900	93.6	388	1 US-07-640-476-7 Sequence 7, Appli
4	1900	93.6	388	1 US-08-112-703-7 Sequence 7, Appli
5	1892.5	93.3	389	1 US-07-840-476-11 Sequence 11, Appl
6	1859	91.6	387	1 US-07-637-870-3 Sequence 3, Appli
7	1859	91.6	387	1 US-07-637-399-9 Sequence 9, Appli
8	1859	91.6	387	1 US-07-640-476-12 Sequence 12, Appl
9	1859	91.6	387	1 US-08-112-703-9 Sequence 9, Appli
10	1734.5	85.5	348	1 US-07-637-399-8 Sequence 8, Appli
11	1734.5	85.5	348	1 US-08-112-703-8 Sequence 8, Appli
12	1729.5	85.2	347	1 US-07-637-870-5 Sequence 5, Appli
13	1729.5	85.2	347	1 US-07-640-476-10 Sequence 10, Appl
14	1353.5	66.7	394	1 US-07-637-870-2 Sequence 2, Appli
15	1353.5	66.7	394	1 US-07-637-399-2 Sequence 2, Appli
16	1353.5	66.7	394	1 US-07-640-476-8 Sequence 8, Appli
17	1353.5	66.7	394	1 US-08-112-703-2 Sequence 2, Appli
18	1351.5	66.6	394	1 US-07-637-870-1 Sequence 1, Appli
19	1351.5	66.6	394	1 US-07-637-399-1 Sequence 1, Appli
20	1351.5	66.6	394	1 US-08-112-703-1 Sequence 1, Appli
21	1347.5	66.4	394	1 US-07-640-476-5 Sequence 5, Appli
22	1344.5	66.3	394	5 5290690-2 Patent No. 5290690
23	1282	63.2	395	1 US-07-637-870-6 Sequence 6, Appli
24	1282	63.2	395	1 US-07-637-399-3 Sequence 3, Appli
25	1282	63.2	395	1 US-07-640-476-9 Sequence 9, Appli
26	1282	63.2	395	1 US-08-112-703-3 Sequence 3, Appli
27	1127	55.5	387	1 US-08-112-630-2 Sequence 2, Appli
28	875	43.1	455	1 US-07-762-681A-2 Sequence 2, Appli
29	320	15.8	444	1 US-08-559-260-2 Sequence 2, Appli

30	317	15.6	444	2	US-08-901-547A-2	Sequence 2, Appli
31	310.5	15.3	439	1	US-07-762-681A-1	Sequence 1, Appli
32	306	15.1	439	1	US-07-637-870-9	Sequence 9, Appli
33	306	15.1	439	1	US-07-637-399-6	Sequence 6, Appli
34	306	15.1	439	1	US-08-112-703-6	Sequence 6, Appli
35	282.5	13.9	440	1	US-07-637-870-7	Sequence 7, Appli
36	282.5	13.9	440	1	US-07-637-399-4	Sequence 4, Appli
37	282.5	13.9	440	1	US-08-112-703-4	Sequence 4, Appli
38	277.5	13.7	440	1	US-07-637-870-8	Sequence 8, Appli
39	277.5	13.7	440	1	US-07-637-399-5	Sequence 5, Appli
40	277.5	13.7	440	1	US-08-112-703-5	Sequence 5, Appli
41	106	5.2	3491	2	US-07-642-734C-2	Sequence 2, Appli
42	95.5	4.7	329	2	US-08-606-143-2	Sequence 2, Appli
43	94.5	4.7	540	4	PCT-US94-06362-3	Sequence 3, Appli
44	93	4.6	540	2	US-08-368-834-20	Sequence 20, Appl
45	93	4.6	541	2	US-08-467-822-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1  
US-07-637-870-4  
Sequence 4, Application US/07637870  
Patent No. 5310665  
GENERAL INFORMATION:  
APPLICANT: Lambier, Anne-Marie  
APPLICANT: Lastiers, Ignace  
APPLICANT: Ouax, Wilhemus J.  
APPLICANT: Van Der Laan, Jan M.  
TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES HAVING  
TITLE OF INVENTION: ALTERED SUBSTRATE SPECIFICITY  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 545 Middlefield Road, Suite 200  
CITY: Menlo Park  
STATE: California  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/637,870  
FILING DATE: 19910104  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murashige, Kate H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 24615-20019.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-813-5600  
TELEFAX: 415-327-2951  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 388 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-07-637-870-4

Query Match 93.6%; Score 1900; DB 1; Length 388;  
Best Local Similarity 92.8%; Pred. No. 2,7e-182;  
Matches 360; Conservative 14; Mismatches 12; Indels 2; Gaps 2;  
QY 1 MNYQPTDEFTGLTWGVGQGRDPFGDATRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59  
Db 1 MSFQPTDEFTGLTWGVGQGRDPFGDATRPALDPVETVQRLAELGAYGVTFHDDDLIP 60

QY	60	FGASDTEREAHVRRFRQALD	QWTVPMATTNLFTHPVFKAGAF	TANDRAVRYALR	KTI	119	
Db	61	FGSSDTERESHIRFRQALD	QWTVPMATTNLFTHPVFKDGF	TANDRDVRYALR	KTI	120	
QY	120	RNTDLAVELGAKVYVANG	REGAESGAANDVRAALDRMKEAF	DLLEGYVTSOGYDIR	FAI	179	
Db	121	GNIDLAAELGAKTVVANG	REGAESGKANDVDRALDRMKEAF	DLLEGYVTAQYDUR	FAI	180	
QY	180	EPXPNPRGDIILLPTIGH	ALAFIERLERPELYGVNPEV	GEQMAGLNFPHGTAQAL	WAGK	239	
Db	181	EPXPNPRGDIILLPTVGH	ALAFIERLERPELYGVNPEV	GEQMAGLNFPHGTAQAL	WAGK	240	
QY	240	LPHIDLNGQSGGIKYDQ	LFRFGAGDLRAAFWLVDL	LESAGWEGPRHDF	FPFPRTEDIDG	WV	299
Db	241	LPHIDLNGQSGGIKYDQ	LFRFGAGDLRAAFWLVDL	LESAGWEGPRHDF	FPFPRTEDIDG	WV	300
QY	300	ASAAGCMRNYLIILKRAA	FRADPEVQEALRAARLDQAL	AEPTAADQLQALLADR	AYEDF	359	
Db	301	ASAAGCMRNYLIILKRAA	FRADPEVQEALRAARLDQAL	AEPTAADQLQALLADR	AYEDF	360	
QY	360	DVD-AAARGMAFERLDQAL	MWDHLLGARG	386			
Db	361	DVDAAAARGMAFERLDQAL	MWDHLLGARG	388			

```

RESULT      2
US-07-637-399-7
; Sequence 7, Application US/07637399
; Patent No. 5340738
; GENERAL INFORMATION:
; APPLICANT: Lambeir, Anne-Marie
; APPLICANT: Lasters, Ignace
; APPLICANT: Mrabet, Nadir
; APPLICANT: Quax, Wilhelmus J.
; APPLICANT: Van Der Laan, Jan M.
; APPLICANT: Misset, Onno
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES WITH AN
; TITLE OF INVENTION: ALTERED PH OPTIMUM
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/637,399
; FILING DATE: 19910104
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 4615-0020.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-637-399-7

```

```

3
RESULT 3
US-07-640-476-7
; Sequence 7, Application US/07640476
; Patent No. 5376536
; GENERAL INFORMATION:
; APPLICANT: QUAX, WILHELMUS
; APPLICANT: LUITEN, RUDOLF G.M.
; APPLICANT: SCHUURHUIZEN, PAUL W.
; APPLICANT: MRABET, NADIR
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASE ENZYMES AND
; TITLE OF INVENTION: THEIR USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/640,476
; FILING DATE: 19910110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kate H. Murashige
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20009.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-7250
; TELEFAX: (415) 327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids

```

TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-640-476-7

Query Match 93.6%; Score 1900; DB 1; Length 388;  
Best Local Similarity 92.8%; Pred. No. 2.7e-182;  
Matches 360; Conservative 14; Mismatches 12; Indels 2; Gaps 2;

QY 1 MNYQPTPEDRFTGLTWGQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59  
DB 1 MSQPTPEDRFTGLTWGQGRDPFGDTRPALDPVETVQRLAELGAYGVTFHDDDLIP 60  
QY 60 FGASDTEREAHVRFKQALDQATGTVPMATTNLFTHPVFKAGAFANDRAVRYALRKT 119  
DB 61 FGSSDTERESHKRFKQALDQATGTVPMATTNLFTHPVFKDGGFTANDRDVRYALRKT 120  
QY 120 RNIDLAELGAKYVAVWGREGAESGAADVRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179  
DB 121 GNIDLAELGAKYVAVWGREGAESGAADVRAALDRMKEAFDILLGEYVTAQGYDLRFAI 180  
QY 180 EKPNEPRGDILLPTTGHALAFIERLERPELYGVNPEVGEQWAGLNFPHGIAQALWAGK 239  
DB 181 EKPNEPRGDILLPTTGHALAFIERLERPELYGVNPEVGEQWAGLNFPHGIAQALWAGK 240  
QY 240 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGVW 299  
DB 241 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGVW 300  
QY 300 ASAAGCMRNYLILKERAAAFRADPEVQEARLRAARLDQALAEPTAADGLQALLADRTAYEDF 359  
DB 301 ASAAGCMRNYLILKERAAAFRADPEVQEARLRAARLDQALAEPTAADGLQALLADRTAYEDF 360  
QY 360 DVD-AAARGMAFERLDQALMDHLLGARG 386  
DB 361 DVDAARGMAFERLDQALMDHLLGARG 388

RESULT 4

US-08-112-703-7  
Sequence 7, Application US/08112703  
Patent No. 5384257

GENERAL INFORMATION:  
APPLICANT: Lambel, Anne-Marie  
APPLICANT: Lasters, Ignace  
APPLICANT: Mrabet, Nadir  
APPLICANT: Quax, Wilhelmus J.  
APPLICANT: Van Der Laan, Jan M.  
APPLICANT: Misset, Onno

TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES WITH AN  
TITLE OF INVENTION: ALTERED PH OPTIMUM

NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/112,703  
FILING DATE: 26-AUG-1993  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Murashige, Kate H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 4615002001

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-813-5600  
TELEFAX: 415-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 388 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-112-703-7

Query Match 93.6%; Score 1900; DB 1; Length 388;  
Best Local Similarity 92.8%; Pred. No. 2.7e-182;  
Matches 360; Conservative 14; Mismatches 12; Indels 2; Gaps 2;

QY 1 MNYQPTPEDRFTGLTWGQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59  
DB 1 MSQPTPEDRFTGLTWGQGRDPFGDTRPALDPVETVQRLAELGAYGVTFHDDDLIP 60  
QY 60 FGASDTEREAHVRFKQALDQATGTVPMATTNLFTHPVFKAGAFANDRAVRYALRKT 119  
DB 61 FGSSDTERESHKRFKQALDQATGTVPMATTNLFTHPVFKDGGFTANDRDVRYALRKT 120  
QY 120 RNIDLAELGAKYVAVWGREGAESGAADVRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179  
DB 121 GNIDLAELGAKYVAVWGREGAESGAADVRAALDRMKEAFDILLGEYVTAQGYDLRFAI 180  
QY 180 EKPNEPRGDILLPTTGHALAFIERLERPELYGVNPEVGEQWAGLNFPHGIAQALWAGK 239  
DB 181 EKPNEPRGDILLPTTGHALAFIERLERPELYGVNPEVGEQWAGLNFPHGIAQALWAGK 240  
QY 240 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGVW 299  
DB 241 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGVW 300  
QY 300 ASAAGCMRNYLILKERAAAFRADPEVQEARLRAARLDQALAEPTAADGLQALLADRTAYEDF 359  
DB 301 ASAAGCMRNYLILKERAAAFRADPEVQEARLRAARLDQALAEPTAADGLQALLADRTAYEDF 360  
QY 360 DVD-AAARGMAFERLDQALMDHLLGARG 386  
DB 361 DVDAARGMAFERLDQALMDHLLGARG 388

RESULT 5

US-07-640-476-11  
Sequence 11, Application US/07640476  
Patent No. 5376536

GENERAL INFORMATION:  
APPLICANT: QUAX, WILHELMUS  
APPLICANT: LUITEN, RUDOLF G.M.  
APPLICANT: SCHUURHUIZEN, PAUL W.  
APPLICANT: MRABET, NADIR

TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASE ENZYMES AND  
TITLE OF INVENTION: THEIR USE

NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 545 Middlefield Road, Suite 200  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/640,476

```

; FILING DATE: 19910110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kate H. Murashige
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20009.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-7250
; TELEFAX: (415) 327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORGANISM: Streptomyces violaceoniger
; STRAIN: CBS 409.73
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..389
; PUBLICATION INFORMATION:
; AUTHORS: Tiraby,
; JOURNAL: Nucleic Acids Res.
; VOLUME: 16
; PAGES: 9337-
; DATE: 1988
;
US-07-640-476-11

```

```

Query Match      93.3%; Score 1892.5; DB 1; Length 389;
Best Local Similarity 92.3%; Pred. No. 1.5e-181;
Matches 359; Conservative 15; Mismatches 12; Indels 3; Gaps 2;

```

```

Qy 1 MNYQPTPEDRTFFGLTWVGWQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MSFQTPEDKFTFGLTWVGWQGRDPFGDTRPALDPVETVQRLAELGAYGVTFHDDDLIP 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 60 FGASDTERAHVKRFRQALDAGTMTVPMTNLFTHPVFKAGFTANDRAVRRYALRKT 119
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 FGSSDTERESHKFRQALDAGTMTVPMTNLFTHPVFKDGGFTANDRDRYALRKT 120
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 120 RNIDLAVELGAKVYVWVGREGAESAAGKDVRAALDRMKEAFDILLGEYVTSOGYDIRFAI 179
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 RNIDLAAELGAKTYVWVGREGAESAAGKDVRAALDRMKEAFDILLGEYVTAQGYDLRFAI 180
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 180 EPKPNPEPGDILLPTIGHALAFIERLERPELYGVNPEVGHQAGLNFPHGIAQALWAGK 239
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 EPKPNPEPGDILLPTVGHALAFIERLERPELYGVNPEVGHQAGLNFPHGIAQALWAGK 240
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 240 LFHIDLNGSQGKIKYDQDLRFAGDRLAAAFWLVLDLESAGWEGPRHFDKPPRTEDIDGVW 299
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 LFHIDLNGSQGKIKYDQDLRFAGDRLAAAFWLVLDLESAGYEGPRHFDKPPRTEDIDGVW 300
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 300 ASAGCMENYLLKERRAAAFRAADPEVQEAALRAALDQLAEPDGLQALLADRTAFEDF 359
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 ASAGCMENYLLKERRAAAFRAADPEVQEAALRAALDQLAEPDGLQALLADRTAFEDF 360
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 360 DVDAARGMA--FERLDQLANDHLLGARG 386
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 DVEAAAAAARAAWFERLDQLANDHLLGARG 389
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 6
US-07-637-870-3
; Sequence 3, Application US/07637870
; Patent No. 5310665
; GENERAL INFORMATION:
; APPLICANT: Lambier, Anne-Marie
; APPLICANT: Lesters, Ignace

```

```

; APPLICANT: Quax, Wilhemus J.
; APPLICANT: Van Der Laan, Jan M.
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES HAVING
; TITLE OF INVENTION: ALTERED SUBSTRATE SPECIFICITY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/637,870
; FILING DATE: 19910104
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20009.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-813-5600
; TELEFAX: 415-327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-07-637-870-3

```

```

Query Match      91.6%; Score 1859; DB 1; Length 387;
Best Local Similarity 91.2%; Pred. No. 3.4e-178;
Matches 353; Conservative 15; Mismatches 17; Indels 2; Gaps 2;

```

```

Qy 1 MNYQPTPEDRTFFGLTWVGWQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MNYQPTPEDRTFFGLTWVGWQGRDPFGDTRPALDPVSRVRLSELGAYGVTFHDDDLIP 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 60 FGASDTERAHVKRFRQALDAGTMTVPMTNLFTHPVFKAGFTANDRAVRRYALRKT 119
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 FGSSDTERESHKFRQALDAGTMTVPMTNLFTHPVFKDGAFTANDRDRYALRKT 120
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 120 RNIDLAVELGAKVYVWVGREGAESAAGKDVRAALDRMKEAFDILLGEYVTSOGYDIRFAI 179
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 RNIDLAVELGASVYVWVGREGAESAAGKDVRAALDRMKEAFDILLGEYVTSOGYDLRFAI 180
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 180 EPKPNPEPGDILLPTIGHALAFIERLERPELYGVNPEVGHQAGLNFPHGIAQALWAGK 239
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 EPKPNPEPGDILLPTVGHALAFIERLERPELYGVNPEVGHQAGLNFPHGIAQALWAGK 240
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 240 LFHIDLNGSQGKIKYDQDLRFAGDRLAAAFWLVLDLESAGWEGPRHFDKPPRTEDIDGVW 299
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 LFHIDLNGSQGKIKYDQDLRFAGDRLAAAFWLVLDLESAGYAGPRHFDKPPRTEDIDGVW 300
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 300 ASAGCMENYLLKERRAAAFRAADPEVQEAALRAALDQLAEPDGLQALLADRTAFEDF 359
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 ASAGCMENYLLKERRAAAFRAADPEVQEAALRAALDQLAEPDGLQALLADRTAFEDF 360
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 360 DVD-AAARGMAFERLDQLANDHLLGAR 385
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 DVDAARGMAFEHLQLANDHLLGAR 387
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 7
US-07-637-399-9

```



Sequence 9, Application US/07637399  
Patent No. 5340738  
GENERAL INFORMATION:  
APPLICANT: Lambair, Anne-Marie  
APPLICANT: Lasters, Ignace  
APPLICANT: Mrabet, Nadir  
APPLICANT: Quax, Wilhelmus J.  
APPLICANT: Van Der Laan, Jan M.  
APPLICANT: Misset, Onno  
TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES WITH AN  
ALTERED PH OPTIMUM  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 545 Middlefield Road, Suite 200  
CITY: Menlo Park  
STATE: California  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/637,399  
FILING DATE: 19910104  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murashige, Kate H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 4615-0020.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-327-7250  
TELEFAX: 415-327-2951  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-07-637-399-9

Query Match 91.6%; Score 1859; DB 1; Length 387;  
Best Local Similarity 91.2%; Pred. No. 3.4e-178;  
Matches 353; Conservative 15; Mismatches 17; Indels 2; Gaps 2;

QY 1 MNYQPTEDRFTFGLTVGQGRDPFGDATRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59  
|||||  
Db 1 MNYQPTEDRFTFGLTVGQGRDPFGDATRPALDPVD-VQRLAELGAYGVTFHDDDLIP 60  
|||||  
QY 60 FGASDTEREHAHVRFQALDQATGTVPMATTNLFTHPVFKAGAFANDRAVRYALRKT 119  
|||||  
Db 61 FGSSDTERESHIRFRQALDQATGKVPMTNLFTHPVFKDGAFTANDRDVRYALRKT 120  
|||||  
QY 120 RNTDLAVELGAKYVYVWVGREGAESGAADVRAALDRMKEAFDLLGLEYVTSQGYDIRFAI 179  
|||||  
Db 121 RNTDLAVELGASVYVWVGREGAESGAADVRAALDRMKEAFDLLGLEYVTEQGYDLKFAI 180  
|||||  
QY 180 EPKPNPRGDIILPTIGHALAFIERLERPELYGNVPVGEQHMAGLNFPHGIAQALWAGK 239  
|||||  
Db 181 EPKPNPRGDIILPTVGHALAFIERLERPELYGNVPVGEQHMAGLNFPHGIAQALWAGK 240  
|||||  
QY 240 LFHIDLNGQSGIKYDQDLRGAGDLRAAFWLVDLLSAGWEGPRHFDKPPRTEDIDGVW 299  
|||||  
Db 241 LFHIDLNGQSGIKYDQDLRGAGDLRAAFWLVDLLERAGYAGPRHFDKPPRTEDIDGVW 300  
|||||  
QY 300 ASAGCMRNYLILKEAARAFRAPEVQEAALRAARLDOLASPTAADGLQALLADRTAYEDF 359  
|||||  
Db 301 ASAGCMRNYLILKDRAAAFRAFPQVQEAALAAARLDLARPTAEDGLAALLADRSAYDTF 360  
|||||

QY 360 DVD-AAARGMAFERLDQALMDHLLGAR 385  
|||||  
Db 361 DVDAAAARGMAFEHLDQALMDHLLGAR 387  
|||||  
RESULT 8  
US-07-640-476-12  
Sequence 12, Application US/07640476  
Patent No. 5376536  
GENERAL INFORMATION:  
APPLICANT: QUAX, WILHELMUS  
APPLICANT: LUITEN, RUDOLF G.M.  
APPLICANT: SCHUURHUIZEN, PAUL W.  
APPLICANT: MRABET, NADIR  
TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASE ENZYMES AND  
THEIR USE  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 545 Middlefield Road, Suite 200  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/640,476  
FILING DATE: 19910110  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kate H. Murashige  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 24615-20009.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 327-7250  
TELEFAX: (415) 327-2951  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Streptomyces violaceoruber  
STRAIN: LMG 7183  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..387  
US-07-640-476-12

Query Match 91.6%; Score 1859; DB 1; Length 387;  
Best Local Similarity 91.2%; Pred. No. 3.4e-178;  
Matches 353; Conservative 15; Mismatches 17; Indels 2; Gaps 2;

QY 1 MNYQPTEDRFTFGLTVGQGRDPFGDATRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59  
|||||  
Db 1 MNYQPTEDRFTFGLTVGQGRDPFGDATRPALDPVD-VQRLAELGAYGVTFHDDDLIP 60  
|||||  
QY 60 FGASDTEREHAHVRFQALDQATGTVPMATTNLFTHPVFKAGAFANDRAVRYALRKT 119  
|||||  
Db 61 FGSSDTERESHIRFRQALDQATGKVPMTNLFTHPVFKDGAFTANDRDVRYALRKT 120  
|||||  
QY 120 RNTDLAVELGAKYVYVWVGREGAESGAADVRAALDRMKEAFDLLGLEYVTSQGYDIRFAI 179  
|||||  
Db 121 RNTDLAVELGASVYVWVGREGAESGAADVRAALDRMKEAFDLLGLEYVTEQGYDLKFAI 180  
|||||



Query Match 85.5%; Score 1734.5; DB 1; Length 348;  
Best Local Similarity 94.0%; Pred. NO. 8.4e-166;  
Matches 327; Conservative 13; Mismatches 7; Indels 1; Gaps 1;

RESULT 12  
US-07-637-870-5  
; Sequence 5, Application US/07637870  
; Patent No. 5310665  
; GENERAL INFORMATION:  
; APPLICANT: Lambier, Anne-Marie  
; APPLICANT: Lasters, Ignace  
; APPLICANT: Quax, Wilhemus J.  
; APPLICANT: Van Der Laan, Jan M.  
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES HAVING  
; TITLE OF INVENTION: ALTERED SUBSTRATE SPECIFICITY  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 545 Middlefield Road, Suite 200  
; CITY: Menlo Park  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk.  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/637,870  
; FILING DATE: 19910104  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murashige, Kate H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 24615-20019.00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-813-5600  
; TELEFAX: 415-327-2951  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 347 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-07-637-870-5

```

; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Streptomyces thermovulgaris
; STRAIN: DSM 40444
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..347
; OTHER INFORMATION: /note= "The C-terminal region of
; OTHER INFORMATION: this protein is not complete."
US-07-640-476-10

Query Match      85.2%; Score 1729.5; DB 1; Length 347;
Best Local Similarity 93.9%; Pred. No. 2.7e-165;
Matches 326; Conservative 13; Mismatches 7; Indels 1; Gaps 1;

Qy 1 MNYQPTPEDRTFFGLTWVGWGRDPFGDTRPALDPV-DVQRLAELGAYGVTFHDDDLIP 59
Db 1 MNYQPTPEDRTFFGLTWVGWGRDPFGDTRPALDPV-DVQRLAELGAYGVTFHDDDLIP 60
Qy 60 FGASDTEREAHVKFRQALDAGTMTVPMTNLTHTPVFKAGFTANDRAVRRYALRRTI 119
Db 61 FGASEAEREAEHVKFRQALDAGTMTVPMTNLTHTPVFKAGFTANDRAVRRYALRRTI 120
Qy 120 RNIDLAVELGAKYVYVANGREGAESAAGKDVRAALDRMKEAFDLLGEYVTSQGYDIRFAI 179
Db 121 RNIDLAVELGARTYVANGREGAESAAGKDVRAALDRMKEAFDLLGEYVTSQGYDIRFAI 180
Qy 180 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQAGLNPFGHIAQALWAGK 239
Db 181 EPKNEPRGDILLPTVGHALAFIERLERPELVGNPEVGHQAGLNPFGHIAQALWAGK 240
Qy 240 LFHIDLNGSQGIKYDQDLRFAGAGDLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVM 299
Db 241 LFHIDLNGSQGIKYDQDLRFAGAGDLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVM 300
Qy 300 ASAAGCMRNYLILKERAFAFRADPEVQALRAARLDQLAEPDAAADGL 346
Db 301 ASAAGCMRNYLILKERSAFAFRADPEVQALRAASRLDQLAQPDAADGL 347

RESULT 13
US-07-640-476-10
; Sequence 10, Application US/07640476
; Patent No. 5376536
; GENERAL INFORMATION:
; APPLICANT: QUAX, WILHELMUS
; APPLICANT: LUITEN, RUDOLF G.M.
; APPLICANT: SCHUURHUIZEN, PAUL W.
; APPLICANT: MRABET, NADIR
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASE ENZYMES AND
; TITLE OF INVENTION: THEIR USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/640,476
; FILING DATE: 19910110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kate H. Murashige
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20009.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-7250
; TELEFAX: (415) 327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Streptomyces thermovulgaris
; STRAIN: DSM 40444
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..347
; OTHER INFORMATION: /note= "The C-terminal region of
; OTHER INFORMATION: this protein is not complete."
US-07-640-476-10

Query Match      85.2%; Score 1729.5; DB 1; Length 347;
Best Local Similarity 93.9%; Pred. No. 2.7e-165;
Matches 326; Conservative 13; Mismatches 7; Indels 1; Gaps 1;

Qy 1 MNYQPTPEDRTFFGLTWVGWGRDPFGDTRPALDPV-DVQRLAELGAYGVTFHDDDLIP 59
Db 1 MNYQPTPEDRTFFGLTWVGWGRDPFGDTRPALDPV-DVQRLAELGAYGVTFHDDDLIP 60
Qy 60 FGASDTEREAHVKFRQALDAGTMTVPMTNLTHTPVFKAGFTANDRAVRRYALRRTI 119
Db 61 FGASEAEREAEHVKFRQALDAGTMTVPMTNLTHTPVFKAGFTANDRAVRRYALRRTI 120
Qy 120 RNIDLAVELGAKYVYVANGREGAESAAGKDVRAALDRMKEAFDLLGEYVTSQGYDIRFAI 179
Db 121 RNIDLAVELGARTYVANGREGAESAAGKDVRAALDRMKEAFDLLGEYVTSQGYDIRFAI 180
Qy 180 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQAGLNPFGHIAQALWAGK 239
Db 181 EPKNEPRGDILLPTVGHALAFIERLERPELVGNPEVGHQAGLNPFGHIAQALWAGK 240
Qy 240 LFHIDLNGSQGIKYDQDLRFAGAGDLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVM 299
Db 241 LFHIDLNGSQGIKYDQDLRFAGAGDLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVM 300
Qy 300 ASAAGCMRNYLILKERAFAFRADPEVQALRAARLDQLAEPDAAADGL 346
Db 301 ASAAGCMRNYLILKERSAFAFRADPEVQALRAASRLDQLAQPDAADGL 347

RESULT 14
US-07-637-870-2
; Sequence 2, Application US/07637870
; Patent No. 5310665
; GENERAL INFORMATION:
; APPLICANT: Lambier, Anne-Marie
; APPLICANT: Lasters, Ignace
; APPLICANT: Quax, Wilhemus J.
; APPLICANT: Van Der Leen, Jan M.
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES HAVING
; TITLE OF INVENTION: ALTERED SUBSTRATE SPECIFICITY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/637,870
; FILING DATE: 19910104
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20019.00
```

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-813-5600  
 TELEFAX: 415-327-2951  
 TELEX: 706141  
 INFORMATION FOR SEQ ID NO: 2:  
 NAME: Murashige, Kate H.  
 REGISTRATION NUMBER: 29,959  
 REFERENCE/DOCKET NUMBER: 4615-0020.00  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 394 amino acids  
 TYPE: AMINO ACID  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-07-637-870-2

Query Match 66.7%; Score 1353.5; DB 1; Length 394;  
 Best Local Similarity 66.0%; Pred. No. 1.4e-127;  
 Matches 260; Conservative 46; Mismatches 79; Indels 9; Gaps 4;

QY 1 MNYQPTPEDRFTFGLTWGQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59  
 DB 1 MSLQATPDDKFSFGLTWGQARDAFGDTRPVLDPIEAVHKLAEIGAYGVTFHDDDLVP 60  
 QY 60 FGADTEREAHKRRFRQALDGMTVPMATTNLFTHPVFKAGAFANDRAVRRYALRKTI 119  
 DB 61 FGADATRGIVAGFSKALDETGLIVPMVTNLFTHPVFKDGGFTSNDRSVRYAIRKVL 120  
 QY 120 RNIDLAELGAKYVAVWGREGAESAKADVRAALDRMKEAFDLLGEYVTSQGYDIRFAI 179  
 DB 121 ROMDLGAELGAKTLVLWGREGAEYDSAKDVGAALDRYREALNLLAQYSEDQGYGLPFAI 180  
 QY 180 EKPNEPRGDILLPTIGHALAFIERLERPELVGNVEVGEQMGAGLNFPHGIAQALWAGK 239  
 DB 181 EKPNEPRGDILLPTAGHAIQVQELERPELVGNVEVGEQMGAGLNFPHGIAQALWAGK 240  
 QY 240 LFHIDLNGSGIKYDQDLRFAGDLRAAFWLVLDLSEA-----GWEGPRHFDKPPRTED 294  
 DB 241 LFHIDLNGSGIKYDQDLRFAGDLRAAFWLVLDLSEA-----GWEGPRHFDKPPRTED 300  
 QY 295 IDGVWASAGCMRNLYILKERAAAFRADPEVQEAALRAALDQAEPTAADG--LQALLAD 352  
 DB 301 FDGVWESAKDNIRMYLLKRAKAFRADPEVQEAALRAALDQAEPTAADG--LQALLAD 360  
 QY 353 RTAYEDFDVDA-AARGMAFERLDOLAMDHLGAR 385  
 DB 361 RSAFEDYDADAVGAKGYGVFKLNQLAIDLHLLGAR 394

RESULT 15

US-07-637-399-2  
 Sequence 2, Application US/07637399  
 Patent No. 5340738

GENERAL INFORMATION:  
 APPLICANT: Lambel, Anne-Marie  
 APPLICANT: Lasters, Ignace  
 APPLICANT: Mrabet, Nadir  
 APPLICANT: Quax, Wilhelms J.  
 APPLICANT: Van Der Laan, Jan M.  
 APPLICANT: Missset, Onno  
 TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASES WITH AN  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morrison & Foerster  
 STREET: 545 Middlefield Road, Suite 200  
 CITY: Menlo Park  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94025

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/637,399  
 FILING DATE: 19910104  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Murashige, Kate H.  
 REGISTRATION NUMBER: 29,959  
 REFERENCE/DOCKET NUMBER: 4615-0020.00  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-327-2951  
 TELEFAX: 415-327-2951  
 TELEX: 706141  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 394 amino acids  
 TYPE: AMINO ACID  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-07-637-399-2

Query Match 66.7%; Score 1353.5; DB 1; Length 394;  
 Best Local Similarity 66.0%; Pred. No. 1.4e-127;  
 Matches 260; Conservative 46; Mismatches 79; Indels 9; Gaps 4;

QY 1 MNYQPTPEDRFTFGLTWGQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59  
 DB 1 MSLQATPDDKFSFGLTWGQARDAFGDTRPVLDPIEAVHKLAEIGAYGVTFHDDDLVP 60  
 QY 60 FGADTEREAHKRRFRQALDGMTVPMATTNLFTHPVFKAGAFANDRAVRRYALRKTI 119  
 DB 61 FGADATRGIVAGFSKALDETGLIVPMVTNLFTHPVFKDGGFTSNDRSVRYAIRKVL 120  
 QY 120 RNIDLAELGAKYVAVWGREGAESAKADVRAALDRMKEAFDLLGEYVTSQGYDIRFAI 179  
 DB 121 ROMDLGAELGAKTLVLWGREGAEYDSAKDVGAALDRYREALNLLAQYSEDQGYGLPFAI 180  
 QY 180 EKPNEPRGDILLPTIGHALAFIERLERPELVGNVEVGEQMGAGLNFPHGIAQALWAGK 239  
 DB 181 EKPNEPRGDILLPTAGHAIQVQELERPELVGNVEVGEQMGAGLNFPHGIAQALWAGK 240  
 QY 240 LFHIDLNGSGIKYDQDLRFAGDLRAAFWLVLDLSEA-----GWEGPRHFDKPPRTED 294  
 DB 241 LFHIDLNGSGIKYDQDLRFAGDLRAAFWLVLDLSEA-----GWEGPRHFDKPPRTED 300  
 QY 295 IDGVWASAGCMRNLYILKERAAAFRADPEVQEAALRAALDQAEPTAADG--LQALLAD 352  
 DB 301 FDGVWESAKDNIRMYLLKRAKAFRADPEVQEAALRAALDQAEPTAADG--LQALLAD 360  
 QY 353 RTAYEDFDVDA-AARGMAFERLDOLAMDHLGAR 385  
 DB 361 RSAFEDYDADAVGAKGYGVFKLNQLAIDLHLLGAR 394

Search completed: March 22, 2000; 23:43:55  
 Job time: 781 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2000, 23:04:53 ; Search time 752.35 Seconds  
(without alignments)  
-6239.449 Million cell updates/sec

Title: US-09-383-318-1  
Perfect score: 1546  
Sequence: 1 cagagcgcttggtgact.....gcgcggtgtgggcgcgtgc 1546

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 821193 seqs, -1518192014 residues  
Total number of hits satisfying chosen parameters: 1642386

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

Result No.	Query Match	Score	Length	DB ID	Description
1	100.0	1546	1546	1	SSPY15518
2	1031.4	66.7	1952	2	AF170068 Streptomyces
3	977.2	63.2	3060	1	STMXYLAB Streptomyces
4	972.6	62.9	1201	1	STMXYLI Streptomyces
5	971	62.8	1202	1	SVXYL Streptomyces
6	970.8	62.8	1522	1	SVXYL Streptomyces
7	952.8	61.6	1164	5	AF10243 S.murinus D
8	947.8	61.3	1164	5	AF10243 S.murinus D
9	935	60.5	1167	5	AF10243 S.murinus D
10	904.6	58.5	2010	5	AF10243 S.murinus D
11	625.2	40.4	1892	1	AF10243 S.murinus D
12	599.6	38.8	1567	5	AF10243 S.murinus D
13	599.6	38.8	1567	5	AF10243 S.murinus D
14	598.8	38.7	1185	5	AF10243 S.murinus D
15	598.4	38.7	1182	5	AF10243 S.murinus D
16	531.4	34.4	1905	1	AF10243 S.murinus D
17	481.8	31.2	1805	1	AF10243 S.murinus D
18	481.8	31.2	1805	5	AF10243 S.murinus D
19	108.6	7.0	41622	2	AF10243 S.murinus D
20	103	6.7	37750	2	AF10243 S.murinus D
21	99.2	6.4	42655	1	AF10243 S.murinus D
22	98.6	6.4	1820	1	AF10243 S.murinus D
23	98.6	6.4	1820	5	AF10243 S.murinus D
24	98.6	6.4	40544	1	AF10243 S.murinus D
25	98	6.3	38404	1	AF10243 S.murinus D
26	97	6.3	1821	5	AF10243 S.murinus D
27	96.8	6.3	182626	40	AF10243 S.murinus D
28	96.8	6.3	192592	40	AF10243 S.murinus D
29	96.2	6.2	24700	1	AF10243 S.murinus D
30	95.2	6.2	150724	10	AF10243 S.murinus D
31	95	6.1	3925	2	AF10243 S.murinus D
32	95	6.1	23990	1	AF10243 S.murinus D
33	94.4	6.1	7412	16	AF10243 S.murinus D
34	94	6.1	19830	1	AF10243 S.murinus D
35	93	6.0	5530	1	AF10243 S.murinus D
36	93	6.0	25306	2	AF10243 S.murinus D
37	92.2	6.0	22449	1	AF10243 S.murinus D
38	92.2	6.0	22449	1	AF10243 S.murinus D
39	91.2	5.9	22115	1	AF10243 S.murinus D
40	91.2	5.9	40909	1	AF10243 S.murinus D
41	89.8	5.8	23730	1	AF10243 S.murinus D
42	89.4	5.8	45396	1	AF10243 S.murinus D
43	89.4	5.8	1593	2	AF10243 S.murinus D
44	89.2	5.8	5800	1	AF10243 S.murinus D
45	88.8	5.7	7836	2	AF10243 S.murinus D

ALIGNMENTS

RESULT 1  
SSPY15518 1546 bp DNA  
LOCUS Streptomyces sp. SK strain xyla gene.  
DEFINITION Y15518  
ACCESSION Y15518  
VERSION Y15518.1 GI:4210848  
KEYWORDS xyla gene.

SSPY15518 1546 bp DNA  
Streptomyces sp. SK strain xyla gene.  
Y15518  
Y15518.1 GI:4210848  
xyla gene.

01-FEB-1999





```
|||||
Db 1501 GGGCCCATGCTGCTGCCGGGCGCGGTGTGGGCGCGTGC 1546

RESULT 2
AF170068
LOCUS AF170068 1952 bp DNA BCT 12-AUG-1999
DEFINITION Streptomyces chibaensis D-xylose isomerase (xylA) gene, complete
cds.
ACCESSION AF170068
VERSION AF170068.1 GI:5731284
KEYWORDS Streptomyces chibaensis.
SOURCE Streptomyces chibaensis.
ORGANISM Streptomyces chibaensis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
1 (bases 1 to 1952)
Joo,G.J., Shin,J.H., Heo,G.Y., Park,H.D. and Rhee,I.K.
Streptomyces chibaensis J-59 xylA
Unpublished
2 (bases 1 to 1952)
Joo,G.J., Shin,J.H., Heo,G.Y., Park,H.D. and Rhee,I.K.
Direct Submission
Submitted (19-JUL-1999) Department of Agricultural Chemistry,
College of Agriculture, Kyungpook National University, Buk-gu,
Sankyuk-Dong, 1372, Taegu 702-701, Korea
Location/Qualifiers
1. 1952
/organism="Streptomyces chibaensis"
/strain="J-59"
/db_xref="taxon:67286"
552..1718
/gene="xylA"
552..1718
/gene="xylA"
/codon_start=1
/transl_table=11
/product="D-xylose isomerase"
/protein_id="A484850.1"
/db_xref="GI:5731285"
/translation="MSYQPTPEDRTFLGTVGWGRDPFGDTRALDPVETVQRLA
ELGAGVTHDDLLPFSSDSTDERESHKRFQALDAGMTVPMTATNLFTHPVKDG
AFTANDRVRYALSKTIRNIDLAELGAKTVVANGREGAESGAKDVRSLDRMKE
AFDLGVEYTSQGLDRIAEKPNRPGDILLPVGHAFIERLERELXYNPEV
GHEQAGVPHGIAQALWAGKFIHDLNGOSIKYDQDLRFAGDLRSFVWLDLLE
SAGYEGPHFDPKPRTELDGWSAAGCMRNILKRAAFRAFADPEVQAALRASR
LDLQAQTAADGLEDLLADRAAFEDFVEAAAARGMAFERLDQAMHLLGARG"
BASE COUNT 286 a 740 c 654 g 272 t
ORIGIN

Query Match 66.7%; Score 1031.4; DB 2; Length 1952;
Best Local Similarity 85.0%; Pred. No. 1.6e-84;
Matches 1228; Conservative 0; Mismatches 176; Indels 41; Gaps 5;

QY 2 acgagcgcttggtgactgggtgagcagtgctccacacgacgagcgagcgacccctcggc 61
|||||
Db 326 acgagcgcttggtgactgggtgagcagtgctccacacgacgagcgagcgacccctcggc 384
|||||
QY 62 tgcgacatcgctctccctcttttcccgctcagggctctgacctgagcgttcacgc 121
|||||
Db 385 tgcTGCATCGGGCTCTCCCTCTCCGCGGCTCCGCGGACGCTGCGGTGTGGTTTCA 444
|||||
QY 122 tatgcggcgcttggtggcccggggtgcgacccggcgccgctttctgctccgcgt 181
|||||
Db 445 AGACACCTTGTCTCGAAGACATCTTG----- 472
|||||
QY 182 tcccttcccgaggaacgctgcggcactactaattgttaatcgccctgacgaaatagtcgc 241
|||||
Db 472 TCTCTTCCAGAGGTGCG-TCGGCATACTAATTTGTAACCGGCATGAGCAATAAGTAGTCGC 530
|||||
QY 242 aagcagcgaagcgcggcgaactgaactaccagccaccccgagacaggttcaccttc 301
|||||
```





Qy	1323	ccgcgtacaggaattcgtgacgcggc-----cgccgcgcgatcgcc-ttcgagc	1376
Db	1087	CCCGCTTCGAGGACTTCACGTGGAGGCGCGCGCGCGCATGGCGTTCGAAC	1146
Qy	1377	gcctcgaccagtgcctatgacacacctggtgggcgccgcggctgacccggcg	1431
Db	1147	GCCTCGACCAGTGGCGATGACCACTGCTGGGCGCGCGGCTGACGGTGCG	1201
RESULT	5		
SVXYL			
LOCUS	SVXYL	1202 bp DNA BCT	18-JUL-1995
DEFINITION	Streptomyces violaceoniger xylA gene for D-xylase-isomerase tetramer alpha 4 ; EC 5.3.1.5).		
ACCESSION	X12816		
VERSION	X12816.1	GI:48004	
KEYWORDS	xylA gene; xylase isomerase.		
SOURCE	Streptomyces violaceoniger.		
ORGANISM	Streptomyces violaceoniger		
REFERENCE	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.		
AUTHORS	1 (bases 1 to 1202)		
TITLE	Direct Submission		
JOURNAL	Submitted (09-SEP-1988) Drocourt D., CAYLA, Avenue de Larrieu,		
REFERENCE	31094 Toulouse Cedex, France		
AUTHORS	2 (bases 1 to 1202)		
TITLE	Drocourt,D., Bejar,S., Calmels,T., Reynes,J.P. and Tiraby,G.		
JOURNAL	Nucleotide sequence of the xylase isomerase gene from Streptomyces violaceoniger		
MEDLINE	Nucleic Acids Res. 16 (19), 9337 (1988)		
FEATURES	Location/Qualifiers 1..1202 /organism="Streptomyces violaceoniger" /strain="CBS 409-73" /db_xref="taxon:1953" 12..17 /note="pot. ribosome binding site" 25..1194 /note="D-xylase isomerase (AA 1 - 389)" /codon_start=1 /transl_table=11 /protein_id="CAA31304.1" /db_xref="GI:48005" /db_xref="Swiss-Prot:P09033" /translation="MSFQPTDFKFTGLTVWGQRDPFGFDATPALDPVETVQRLA ELGAVGYTFHDDLLIPGSDSTERSHIKRFOALDATGMVPATNLFHPVKPKK GFPTANDRVRYALKRKIRNIDLAELGAKTYVANGREGAESGAKDVRAIDRMKE AFDLGVYTAQQGDLPFALEPKNEPRGDIILLPTVGHALAFIERLERPELYGNVPE FHQMAGLNFPFGIAQLWAGKLPHIDLNGSGIKYQDRLFAGADLRAAFVLVDLE SAGVEGRHFDFKPRTDEFGYWASGCMRNNTLIILKRAAARADPEVGEALRAAL LDQAOPATAADGLEALLADRTPAFDFVAAAAAAMPFERLDQLAMDHLLGARG"		
RBS			
CDS			
BASE COUNT	187 a	454 c	403 g
ORIGIN	158 t		
Query Match	62.8%	Score 971;	DB 1; Length 1202;
Best Local Similarity	90.0%	Pred. No. 4.7e-79;	
Matches 1076;	Conservative 0;	Mismatches 110;	Indels 9; Gaps 3
Qy	246	gagcaaggagccgcggcatgaactaccagccccccagagacagggttcaacttcggcc	305
Db	8	GAAACAAGGAGCAGGCCCATGAGCTTCCAGCCCCACCCCAGGACAAAGTTACCTTCGGTC	67
Qy	306	tgttgaccgtcgctggcacggggcgggacccttcgcgcgacccacgcgtcccgccctcg	365
Db	68	TGTGACCGTTCGGCTGGCAGGGAAGGGACCCGCTTCGGGACGCCACCCGCTGCCTCG	127
Qy	366	accgggttgcg---cgtgcagcgctggccggaactggggcgcttacgagtgaccttcacg	422
Db	128	ACCGGTCGACACCGTGCAGGCGCTTGGCGGAGCTGGGGCCCTACGAGTAGACTTCCACG	187

```

DEFINITION xylose isomerase [Streptomyces diastaticus, No. 7 M1033, Genomic,
1522 nt].
ACCESSION S73809
VERSION S73809.1 GI:786339
KEYWORDS .
SOURCE Streptomyces diastaticus No. 7 M1033.
ORGANISM Streptomyces diastaticus
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
1 (bases 1 to 1522)
REFERENCE Wang,Y., Huang,Z., Dai,X., Liu,J., Cui,T., Niu,L., Wang,C. and
XU,X.
AUTHORS The sequence of xylose isomerase gene from Streptomyces diastaticus
No. 7 M1033
JOURNAL Chin. J. Biotechnol. 10 (2), 97-103 (1994)
MEDLINE 95101954
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 15677] from the original journal article.
This sequence comes from Fig. 2.
FEATURES             location/qualifiers
     source           1..1522
                     /organism="Streptomyces diastaticus"
                     /db_xref="taxon:1956"
     gene            193..1359
                     /gene="xylose isomerase, XI"
     CDS              193..1359
                     /gene="xylose isomerase, XI"
                     /note="Method: conceptual translation with partial peptide
sequencing. This sequence comes from Fig. 2; XI"
                     /codon_start=1
                     /product="xylose isomerase"
                     /protein_id="AAB32873.1"
                     /db_xref="GI:786340"
                     /translation="MSYQPTPEKFTFGLMTVGWQRDPFGDATRGALDPAESVRRLLA
ELGAFNDVFPHDDLIIPFGADTSRAEHIKRFROLDGETGMKVPMATNLTPTHVPFKDG
GFYANDRVRYAVRKTINIDLAVELGAOTYVANGREGAESAAKDVRVALDRDKGE
AFILLEGVYSQSVDPPFALEPRNPFRGBILLPTIGHALAFIDGLERPELYGNVEPY
SMOAGNLNPFHIGIALWAKGLIFIDLNGSGIKYQQDLRFPGFDIAAAFVLIVLILE
SHQEYGRPHDFPFPTEDEDGVMAASGCMRNYLIIKERAAFFRADPVEQEARLAAR
LDELAOPTAGDGLQLLPDSAFEDEPDAAAAAGMAFERLDQLAMDHLILGARG"
```

```

VERSION      A10243.1  GI:490202
KEYWORDS     glucose isomerase.
SOURCE       Streptomyces murinus.
ORGANISM     Streptomyces murinus
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
REFERENCE    1 (bases 1 to 1164)
AUTHORS      Luiten,R.G.M., Quax,W.J., Schuurhuizen,P.W. and Mrabet,N.
TITLE        Novel glucose isomerase enzymes and their use
JOURNAL      Patent: EP 0351029-A 16 17-JAN-1990;
              GISTR-BROCADES N.V.; PLANT GENETIC SYSTEMS, N.V
FEATURES     Location/Qualifiers
              1..1164
                /organism="Streptomyces murinus"
                /db_xref="taxon:33900"
                1..>1164
                  /codon_start=1
                  /transl_table=11
                  /product="glucose isomerase"
                  /protein_id="CAA0085.1"
                  /db_xref="GI:490203"
                  /db_xref="SWISS-PROT:P37031"
                  /translation="MSFQPTEDRFGLWTVGNGRDPFGDTRPALDPVETVORLA
ELGAYGVTFHDDIIPFSSDTERESHKRFQALDADATGMTPMATTLFTHPVFKDG
AFTDLNDRVRYALRKTGIDLAELGAKTYVAMGGREGASGAKDRPDLDRMKE
AFDILLGEVYTAQGLDRLRAIPKNEPRGDIITLPTVHALAFIERLEPELYGVNPEV
GHEOMAGLNFPHGTAQALWAGKLPKHLIDNGOSGKYDODLRFAGADLPAAPFLVDLLE
TAGYEGPFFKPEKPRTEDEFGVWASAGCMRNVLILKDRAAAFRADEPVEQALRAAR
LDQAQPTAAFGDALLADRAAFEDFDVAAAAGMAFEHLQDLANDHLLGARG"
BASE COUNT  178 a 445 c 382 g 159 t
ORIGIN

Query Match      61.6%; Score 952.8; DB 5; Length 1164;
Best Local Similarity 89.9%; Pred. No. 2e-77;
Matches 1046; Conservative 0; Mismatches 112; Indels 6; Gaps 2;

QY 263 atgaactacagccaccccccagacagcttcacctcgccgtggaccgtcgctg 322
DB 1 ATGAGCTTCACGCCACCCCGGAGGACAGTTCACCTTCGGTCTGTGGACCTCGCTGG 60
QY 323 cggggcgggacccttcggagccagcgctcccgccctcgaccggtcga---cgtg 379
DB 61 CAGGGAAGGACCCGTTCCGGCAGCGCACCCGCCCGCCCTCGACCCGCTGAGACGGTG 120
QY 380 cagcggtcgccgaactggcgccctacagtgagtgacgttcacccagcagacgtgaccc 439
DB 121 CAGCGCTGGCGAGCTGGCGCCTACGGAGTGACCTTCCACGACGACACCTGATCCCC 180
QY 440 ttccggggtccgacacccagcagcgagcgccagtcacgtcagcgctcgac 499
DB 181 TTCGGGTCTCGACACCCAGGCGGAGTGCACATCAACGGTTCCGGCAGGCCCTGGAC 240
QY 500 gcgacgggcatgacggttcggtggtgacccacccaccccttcacccaccccttcacg 559
DB 241 GCCACGGCATGACGGTGCCATGGCCACCAACCACTTCCACCCACCCCTCTCAAG 300
QY 560 gaggcggttcacgcgcacacccagcgagtgccgttacgctcagcgcgcgcacac 619
DB 301 GACGGCGGCTTCACCGCCAAACACCGCAGTACGCGGTACGCGGTGCGCAGACGATC 360
QY 620 cgaacatcgatctcggtcgagctggcgccaggttcagtcgctggcgccgcgcg 679
DB 361 GGCACATCGACTGCGGCGCCAACTGGGTGCGACACGATGTCGCCCTGGCGGCGCGT 420
QY 680 gggggtcggttcgggttcggccaaagagctgctgctggccctggacgcgcatgaaggag 739
DB 421 GAGGGCGCGAGTCCGGTGGCGCCAGGACGTGCGGACGCGCCCTCGACCGCATGAAGGAG 480
QY 740 gcttcgacactctcggtcgagtagctcaactcgacgggttagacatccggttcgcatc 799
DB 481 GCGTTCGACCTCTCGCGCAGTAGTCACGCGCCACGGGTACGACCTCGCTTCGCCATC 540

```

FT /organism='Streptomyces griseofuscus' FT CDS  
1. .1164 /product='glucose isomerase'.  
Location/Qualifiers  
1. .1164 /organism='Streptomyces rochei'  
/db\_xref='taxon:1928' 155 t  
BASE COUNT 181 a 441 c 387 g  
ORIGIN  
Query Match 61.3%; Score 947.8; DB 5; Length 1164;  
Best Local Similarity 90.1%; Pred. No. 5.5e-77;  
Matches 1051; Conservative 0; Mismatches 107; Indels 9; Gaps 3;  
QY 263 atgaactaccagcccccagagagaggttcaaccttgcgcttgagacgtcggtg 322  
DB 1 ATGAGCTTCCAGCCACCCCGAGGACAGTTACCTTCGCTCTGGACCGTGGCTG 60  
QY 323 caggggcgagacccttcggcgagcgaagcggtcccgccctcgaccgggtcga---cggtg 379  
DB 61 CAGGGAAGGAGACCGTTGCGGACGCGCACCCGCCCTGGCTCGACCGGTGAGACCGTG 120  
QY 380 cagcgctggcggaactggcgctcagagtgacattccacagacagacattgatcccc 439  
DB 121 CAGCGCTGGCGAGTGGCGGCTACGGAGTACCTTCCACGACGACGACCTGAATCCC 180  
QY 440 ttccggggttcacacaccagcgagcgacgtcgaagcggttcggttcggttcgac 499  
DB 181 TTCGGGTGCTCCACACCGGAGCGGAGTGCACATCAGCGGTTCGCGAGGCGCTGGAC 240  
QY 500 ggagcgggcatgaccttccgtatggccaccacaaactcttccaccaccccgcttcaag 559  
DB 241 GCACACTGGCATGACGTTGCGGTGCGGACGACGACGACGACGACGACGACGACG 300  
QY 560 gcaaggcggttaccgccaacgacgagcgagtcgagcggttcggttcggttcgacgacatc 619  
DB 301 ---GACCGTTTACCCGCCAACGACGCGGAGCTGCGCGCTACCGCGTGCACGACGATC 357  
QY 620 cggaaatcatctcgggttcgagtcggcgacgaaggtctacgtcgctggcgcgccgc 679  
DB 358 CGCAACATCGACCTGCGGCGGAGTGGGCGCCAAAGACGTACGTGCGCTGGCGGCGCGT 417  
QY 680 gaggcgcgagtcggttcggttcggttcggttcggttcggttcggttcggttcggttcggttc 739  
DB 418 GAGGCGCGGAGTCCGCTGCGGCGGAGGAGTGGCGGACGCGGCGGACGCGGCGGCGGAG 477  
QY 740 gcttcgacctgtcggcgagtagtcacctcgcaggttcgagacatccggttcgacctc 799  
DB 478 GCGTTGACCTCTCGGCGAGTACGTACCGCGCGGAGGCTACGACCTCCGCTTCGCGCATC 537  
QY 800 gaggccaaagcgaacgagcgcgcgacatctcgtgccaccatcgcccgcgctc 859  
DB 538 GAGCCCAAGCCCAAGGACCCCGGCGGACATCTCTGCGCCACCGTCTGGCGGCGGCGGAG 597  
QY 860 gcttcacatcagcgctggagcgcccgagctgtacggtgtcaaccccgaggtggggccac 919  
DB 598 GCCTTCATCAGCGCTTGAGCGCGCGGAGCTGTACGGCGTCAACCGGAGGTTCGGCCAC 657  
QY 920 gacagatggcggttcgagcttcgacgagcgatcgcgagctctgtggcgggcggaag 979  
DB 658 GAGCAGATGGCGGCTGAATTCGCGACGCGGATCGCGAGGCGCTGTGGGCGGCGGAG 717  
QY 980 ctcttcacatcagctcaagcgccagtcgacatcaagtcacacagacacacacacacacac 1039  
DB 718 CTCTTCACATCAGCTCAACGCGCGAGTCCGGCATCAAGTACGACGAGGACCTTCGGGTTT 777  
QY 1040 ggcgcggttgacctgctgctgcttctgctgctgctgctgctgctgctgctgctgctgctgctg 1099  
DB 778 GGGCGCGGCGGAGCTGGCGGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837  
QY 1100 gagggtccgcgaccttcgacctcaagcccccgcgagacatcgacgagcggtg 1159

DB 838 GAGGGCCCGCGGACACTTCGACTTCAAGCGCGCGGACCGGACGACGACGACGACGACG 897  
QY 1160 gcttcgcgcgcggtgcatgagcaactacacacacacacacacacacacacacacacacacac 1219  
DB 898 GCCTCGCGCGAGGCTGCATGCGCACTACCTGATCTTCAAGGAGCGCGGCGCGGCTTC 957  
QY 1220 cgtcgccagcccgaggttcagagggccctcgctgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1279  
DB 958 CGCGCGACCGCGAGGTGACGAGGAGGCGCTGCGCGCGCGCTGACACGACGCTGGCCAG 1017  
QY 1280 ccacacgcgcgacgacctcagagccctcgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1339  
DB 1018 CCACCGCGCGGACGCGCTTGGAGGCGCTGCTCGCGACGCGACGCGCTTCGAGGACGCTTC 1077  
QY 1340 gacgtggaacgcg---gcccgcgcgcgatggccttcgagcgcttcgagcgcttcgagcgcttc 1396  
DB 1078 GAGGTGAGGCGCGCGCGCGCGCGCATGCGTTCGACGCGCTCGACGACGCTGGCGATG 1137  
QY 1397 gaccacctgctggcgcgcgcgcgcgctga 1423  
DB 1138 GACCACCTGCTGGGCGCGCGGCTGA 1164  
RESULT 9  
LOCUS I09223 1167 bp PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent WO 8901520.  
ACCESSION I09223  
VERSION I09223.1 GI:588068  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1167)  
AUTHORS Drummond, R.J., Bloch, W., Matthews, B.W., Toy, P.L. and Nicholson, H.H.  
TITLE PROCARYOTIC XULOSE ISOMERASE MUTEINS AND METHOD TO INCREASE PROTEIN  
STABILITY  
JOURNAL Patent: WO 8901520-A 1 23-FEB-1989;  
FEATURES  
Location/Qualifiers  
1. .1167 /organism='unknown'  
BASE COUNT 180 a 426 c 405 g 156 t  
ORIGIN  
Query Match 60.5%; Score 935; DB 5; Length 1167;  
Best Local Similarity 88.8%; Pred. No. 7.6e-76;  
Matches 1036; Conservative 0; Mismatches 125; Indels 6; Gaps 2;  
QY 263 atgaactaccagcccccagagagaggttcaaccttgcgcttgagacgtcggtg 322  
DB 1 ATGAATACACGACCCACCCCGAGGACAGGTTACCTTCGAGACTGTGGACCGTGGCTGG 60  
QY 323 caggcgcgagacccttcggcgagcgaagcggttcgcgccctcgaccgggtcga---cggtg 379  
DB 61 CAGGGAAGGAGACCGTTGCGGACGCGCACCCGCCCTGGCTCGACCGGTGAGACCGTG 120  
QY 380 cagcgctggcggaactggcgctcagagtgacattccacagacacacacacacacacacacacac 439  
DB 121 CGCGCGTGGCGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180  
QY 440 ttccggggttcgacacaccagcgagcgaggttcggttcggttcggttcggttcggttcggttcg 499  
DB 181 TTCGGGTCCACGACGACGAGCGGAGGAGGACGACGACGACGACGACGACGACGACGACG 240  
QY 500 gcaaggcggttaccgccaacgacgagcgagtcgagcggttcggttcggttcggttcggttcggttc 559  
DB 241 GACACCGGATGAAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
QY 560 gcaaggcggttaccgccaacgacgagcgagtcgagcggttcggttcggttcggttcggttcggttc 619  
DB 301 GAGCGGCGGCTTACCGCGCAACGACGCGGAGCTGCGCGGCTACGCGGCTGCGCGGAGCATC 360



QY 620 cggacacatgatctcgctgagctggcgcccaagggtctacgtgcctggcgccgcgc 679  
 Db 361 CGAACATCGACCTCGCGTGCAGCTCGCGCGGAGACCTATGTGGCTGGCGCGCGC 420  
 QY 680 gaggcgaggagtcggtgctgcgcgaagagctgctgctggccctgagcagcatgaaggag 739  
 Db 421 GAGGTCGCGAGTGGGTGGCCCAAGGACGTGCGGGGACGCGCTCGACCGCATGAAGGAG 480  
 QY 740 gcttcgacctgctgctgagtagctacgtcactcgcgagggctagacatcggttcccatc 799  
 Db 481 GCCTTCGACCTGCTCGGCGAGTACGTACCTCCAGGGCTAGACATCGCTTCGCCATC 540  
 QY 800 gagcccaagcgagcagcggcgccgagcactcctgctgcccaccatggcgccacgcctc 859  
 Db 541 GAGCCCAAGCGACGAGCGCGCGGCGACATCTCTCTCCCGACCGTGGCGCGCGCTG 600  
 QY 860 gcttcacatgagcctgagcggcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 919  
 Db 601 CGCTTCATCGAGCGCTGGAGCGACCGAGCTGTACGGCTGTAACCCCGAGTCCGCCAC 660  
 QY 920 gagcagatggcgccgctgaactccgcacgcgcgcgcgcgcgcgcgcgcgcgcgcgc 979  
 Db 661 GAGCAGATGGCGGCTGAATCTCCGCGACGCATCGCGCAGCGCTGTGGCGCGGCAAG 720  
 QY 980 ctctccacatgacactcaagggccagtcggcgacatcctgctgcccaccatggccttc 1039  
 Db 721 CTGTTCACATCGACCTCAACCGCCAGAACGCGCATCAAGTACGACGACCTCGCTTC 780  
 QY 1040 ggcgcggtgacctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1099  
 Db 781 GCGCGGCGGACCTCGCGGCGCGCTTCTGGCTGTGGTGGACCTCTGAGTTCGGCGCTAC 840  
 QY 1100 gaggttcgcgcgcacttcgactcgaagccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1159  
 Db 841 AGCGCGCGCGGCACTTCGACTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
 QY 1160 gcttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1219  
 Db 901 GCTCGCGCGCGCGCTGCATGGGCAACTACCTGATCTCAAGGAGCGTGGCGCGCTTC 960  
 QY 1220 cgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1279  
 Db 961 CGCGCGACCGCGAGGTGCAGAGCGCTGCGCGCGCTCCCGTCTGAGAGTGGCGCGG 1020  
 QY 1280 ccacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1339  
 Db 1021 CCCACGGCGCGCGAGGTTCGAGCGCGCTGCTGAGCGCGCTGCGCTTCGAGAGTTC 1080  
 QY 1340 gacgtggacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1396  
 Db 1081 GACGTGCGAGCGCGCGCGCGCGCTGGATGGCTTCGAGCGCGCTGGACCGCTGGCGATG 1140  
 QY 1397 gacacactgctggcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1423  
 Db 1141 GACCACCTGCTGGCGCGCGCGCGCTGA 1167

RESULT 10  
 LOCUS E01992 E01992 2010 bp DNA PAT 29-SEP-1997  
 DEFINITION DNA sequence coding for glucose isomerase.  
 ACCESSION E01992  
 VERSION E01992.1 GI:2170240  
 KEYWORDS JP 1989137979-A/1.  
 SOURCE Streptomyces rochei.  
 ORGANISM Streptomyces rochei  
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 REFERENCE 1 (bases 1 to 2010)  
 AUTHORS Fukazawa, C. and Kainuma, K.  
 TITLE GLUCOSE ISOMERASE GENE, RECOMBINANT HAVING SAID GENE AND  
 MICROORGANISM CONTAINING SAID RECOMBINANT  
 JOURNAL Patent: JP 1989137979-A 1 30-MAY-1989;

NATL FOOD RES INST  
 OS Streptomyces griseofuscus  
 PN JP 1989137979-A/1  
 PD 30-MAY-1989  
 PF 24-NOV-1987 JP 1987295739  
 PI FUKAZAWA CHIKAFUSA, KAINUMA KEIJI  
 PC C12N15/00, C12N1/16, C12N9/92, (C12N9/92, C12N1:19), PC  
 (C12N9/92,  
 PC C12N1:125), (C12N9/92, C12N1:865);  
 CC strandedness: Double;  
 CC topology: linear;  
 CC hypothetical: No;  
 CC anti-sense: No;  
 CC \*source: strain=S-41;  
 FH Key Location/Qualifiers  
 FT 5'UTR <1..665  
 FT CDS 666..1850  
 FT /product='glucose isomerase'  
 FT 1851..>2010.  
 FT 3'UTR

FEATURES  
 source  
 1..2010  
 /organism='Streptomyces rochei'  
 /db\_xref='taxon:1928'  
 BASE COUNT 292 a 757 c 673 g 288 t  
 ORIGIN

Query Match 58.5%; Score 904.6; DB 5; Length 2010;  
 Best Local Similarity 81.0%; Pred. No. 3.1e-73;  
 Matches 1226; Conservative 0; Mismatches 264; Indels 23; Gaps 14;

QY 2 acgagcgcttggtgactggtgacgagtcacacacgagcagcagcgacccctcgcc 61  
 Db 407 ACCAGGGCTTCGTGGACTGGGTGACGTGCCATCTCCGACGACCA-CGGGGCCCTCCG 465  
 QY 62 tctgacatcgcttccctcttttcccgctcagggtctctgacctgacgtcgcttcacgc 121  
 Db 466 TTCGACAT-GGGTCTCTCCTCTTCGCGGCTCACGCCCATGTGCTGTGGGGCGCTCC 524  
 QY 122 tatgcggcgctgtggcccggtgagcccggtgagcccgccggttctgtgttcggt 181  
 Db 525 GAGGCCCGGAGCGCTCGTGGAGACATCTGTGAAGACACACCTTTGTGGTGAAGACTA 584  
 QY 182 tccctccacgagcgctcg---gcatactaatgtgaaatccctgacgaaatagt 238  
 Db 585 CTTGTCTCTCCAGAGCGGTGGTGCATATAATTTGTAAA-CGCCATGAGGAATAGT 643  
 QY 239 gcaagcgagcaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 298  
 Db 644 CGTACG-GAACAAGAGCAGC-GCATGAGTTCACAGCCACCCCGAGGACAAAGTTCACC 701  
 QY 299 ttcggcctgtgagcgcgtggtgagggcgagcccttcggcgagcgcgcgcgcgcgcgc 358  
 Db 702 TTCGCTGTGTGACCTGCGTGGAGGAAAGGACCGCTTCGGCGACGCCACCGCCCT 761  
 QY 359 gccctcgaccgcgtcga---cgtgcagcgctggtgcgcgcgcgcgcgcgcgcgcgcgcgc 415  
 Db 762 GGCCTCGACCGGTGACACCGTGCAGCGCTGGCGCGCTGGCGCGCTACGAGATGACC 821  
 QY 416 ttccacgacgacgactgatcccttcggggcggtccgcacacgagcgcgcgcgcgcgcgc 475  
 Db 822 TTCCACGACGACGACCTGAATCCCTTCGGTGTGTCGACACCGAGCGGAGTCCGACATC 881  
 QY 476 aagcgttcctcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 535  
 Db 882 AAGCGTTTCGCCAGCGCTGGACGCCATGAGCTGAGCGTCCGATGGCCACCGACGAAAC 941  
 QY 536 ctcttcaccccccgcgtttcaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 595  
 Db 942 CTCCTCACCCACCGCTCTTCAAG---GACCGTTTCACCGCCCAAGACCGCGACGTGGC 998  
 QY 596 cgttacgccttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 655



Db	999	GCCTAGCGTGGCCAGAGATCCGCAACATCGACTGGCGCCGAGTGCGCGCCAG	1058
Qy	656	gtctagctgcctgggcccgcgagggcgagtcctccggtgcccgaagagcgtgcgt	715
Db	1059	ACGTAGCTCGCTGGCGCGCGCTGAGGCGCGAGTCCGCTGGCGCCAGGAGCTGCGC	1118
Qy	716	ggcgccctggaccgcatgaagggagccttcagacctcctgcgcgagtcacgtcaoctgcag	775
Db	1119	GAGCGCCTCGACCGCATGAAGGAGGGTTGCACTCTCTCGCGAGTACGTACCGCCAG	1178
Qy	776	ggctagacatccggttcgcatcgagcccaagccgaagcgcgcgcgcgcacatcctg	835
Db	1179	GGCTAGACTTCGCTTCGCCATCGAGCCCAACGCCAACAGAGCCCGCGCGGACATCCTC	1238
Qy	836	ctgcccaccatcgccacgcgtcgcttcctcatcgagcgcctggagcgcgcgcgcgcgtac	895
Db	1239	CTGCCACCGTCGGCCACGCCCTTGCCTTCATCGAGCGCTGGAGCGCGCGAGCTGTAC	1298
Qy	896	ggtgtcaaccccgaggtggccacagagcagatggcgcgccctgaactcccgacgcgcac	955
Db	1299	GGCGTCAACCCGAGGTGCGGCCACGACGAGATGCCGCGCTGAACCTCCCGCACGCGATC	1358
Qy	956	gqcgagctctgtggcggaagctcttcacatcgacctcaacgcgcgcgcgcgcgc	1015
Db	1359	GCACAGCCCTGTGGCGGCGCAAGCTTCCACATCGACCTCAACGGCCAGTCCGGCATC	1418
Qy	1016	aagtagaccagagacatcgcttcgctggcgccgctgacctgcgcgcgcgcgcgcgcgc	1075
Db	1419	AAGTAGACACGAGACTGCGG- -ITCGCGGCGCGACCTCGCGGCGGCTTCTGG-TGGTC	1475
Qy	1076	gaactctggagagcgccgctgggaggttcgcgcgaacttcgaactccaagcccgcg	1135
Db	1476	GACCTCTCGAGAGCGCGGTTACGAGGCGCGCGACCTTCGACTCAAGCGCGCGCGG	1535
Qy	1136	accgagacatcgacgcgtgtggcctccgcgcgcgcgcgcgcgcgcgcgcgcgcgc	1195
Db	1536	ACCGAGACTTCACGCGGTGTGGGCTCGCGCGAGGGCTGCATGCGCACTACCTGATC	1595
Qy	1196	ctgaagagcgccgcgcgccttcctgctgcgcgcgcgcgcgcgcgcgcgcgcgcgc	1252
Db	1596	CTCAAGCAGC-CGCGCGCGCTTCCGCGCGCGACCGGAGGTGCAGGCGCGCGCTCGGC	1654
Qy	1253	ggcgccgctgacacagctcgccgagcccaacgcgc-ggcgacgcgcgcgcgcgcgcgc	1311
Db	1655	GCGCGCGCTGAGACAGCTGGCGCCAGCCGACCGCTGGCGGAGCGCTGGAGCGCCTGCT	1714
Qy	1312	ggcgaccgacacgcgtacgagacttcgactgcgcgcgcgcgcgcgcgcgcgcgcgc	1370
Db	1715	CGCGACCGACCGCTTCGAGACTTCGAGCTGGAGGCGCGCGCGCGCGCATGGT	1774
Qy	1371	tcgagccttcgaccagcttcgcatggaccacatgctggcgcccgcgctgaacggcg	1430
Db	1775	TCAAGCGCTCGACAGCTGGCGATGGACACCTGCTGGCGCGCGCGCGCTGACGCTCG	1834
Qy	1431	gacgaggggtagcgcgcgtcgatctcctcgctcgatcgatgagggggtgctggcggtc	1490
Db	1835	GCCAGGAAGCGGTGACGGAAGAACCGGGGACGCGCGCTGCGCGCGCGCTGCTCC	1894
Qy	1491	gagcgcgccgcgc 1503	
Db	1895	CGGCGCTTCGC 1907	
RESULT	11		
AMXXYLI			
LOCUS	1892 bp	DNA	BCT 21-MAR-1996
DEFINITION	Actinoplanes sp. ATCC 31351 xylose isomerase gene, complete cds;		
ACCESSION	M15050.1		
VERSION	GI:141993		
KEYWORDS	xylose isomerase; xylose kinase.		
SOURCE	Actinoplanes ATCC31351 DNA.		
ORGANISM	Actinoplanes ATCC31351		
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
Actinomycetales; Micromonosporineae; Micromonosporaceae;			
Actinoplanes.			
REFERENCE	1 (bases 1 to 1892)		
AUTHORS	Saari,G.C., Kumar,A.A., Kawasaki,G.H., Insley,M.Y. and O'Hara,P.J.		
TITLE	Sequence of the Ampullariella sp. strain 3876 gene coding for xylose isomerase		
JOURNAL	J. Bacteriol. 169 (2), 612-618 (1987)		
MEDLINE	87109047		
FEATURES	Location/Qualifiers		
source	1..1892		
	/organism="Actinoplanes ATCC31351"		
	/strain="Dow Chemical Co. 3876"		
	/db_xref="taxon:1872"		
	220..1404		
CDS	/codon_start=1		
	/transl_table=11		
	/product="xylose isomerase"		
	/protein_id="AA092578.1"		
	/db_xref="GI:141994"		
	/translation="MSLQATPDKESFGLWTVGQARDAFGDATRPVLDPIBAVHKLA		
	EIGAYGVTHDDLVFFGADATRDGIVAGFSKALDETLGLIVPMTNLTFTVPRKDG		
	GFTSNDRSVRRVARIKRLQMDLGAELGAKTLVLMGREGAIDYDAKQVGAALDRRE		
	ALNLAQYSEDQGYGLPFAIEPKNEPRGDIILLTAGHAIAFVQELERPELSIVRET		
	GHEQSNLNFQTGIAQALWHKLFHIDLNGOHGPKEDODLVFGHGLDILNFAISIVLLE		
	NGPDGPAYDGRPLNFHDYRPSRTEDFDGVWESAKNIRMYLLKRAKAFRAPEVOAA		
	LAESKVDLRLTLPNGETIADLADRSAFEDIDADAVGAKGYGVFKLNQLAIDHLLG		
	AR"		
	1407..>1892		
CDS	/note="putative"		
	/codon_start=1		
	/transl_table=11		
	/product="xylose kinase"		
	/protein_id="AA092579.1"		
	/db_xref="GI:141995"		
	/translation="NALVAGIGOLDVAVQGHSAIATPRPATGPAGPAHSGRHOYDPDA		
	HWARATGDSREAGIGRLAAASVAGQQRHGALLESAVTVPRLALLNDTRPQAAA		
	DLIQELGGADKWAENAVGVIPVASLITLNSGLARHEPANAAKVAICLPHDLWTWKL		
	GS"		
BASE COUNT	319 a 671 c 643 g 259 t		
ORIGIN	1033 bp upstream of the Sali site.		
Query Match	40.4%; Score 625.2; DB 1; Length 1892;		
Best Local Similarity	73.3%; Pred. No. 2.4e-48;		
Matches	865; Conservative 0; Mismatches 288; Indels 27; Gaps 4;		
Qy	271	ccagccaccccgagagcagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	330
Db	228	CCAGGCCACACCCGATGACAAAGTTCCTTCGCTCTCTGGACCTCGCTGGCAGCGCGC	287
Qy	331	ggaccccttcggcgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	387
Db	288	TGACGCGTTCGCTGACGCCACCGTCCGGTCTCGACCCGATCGAGGCCGTGCACAACT	347
Qy	388	ggcgacaccccgagcagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	447
Db	348	GGCGGAGATCGCGCGTACGCGCTCACGTTCACGACGACGACCTGGTGGCTGGCGCGC	407
Qy	448	gtccgacacccgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	507
Db	408	CGACGCGCGACCGGACGGGATCGTCGCGGGGTCTCCAAAGCGCTCCACGAGACCGG	467
Qy	508	catgacgcttcgatggccaccaccacacacacacacacacacacacacacacacacacac	567
Db	468	CCTGATGCTCCGATGGTCCACCAACCTGTTACCCACCCCGTGTTCGAAGCAGCGCG	527
Qy	568	gttcacgcacacgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	627
Db	528	CTTCACGACGACGACGCGGATGCGCGGTATGCGATCCGCAAGGTGTGTGCGCCAGAT	587
Qy	628	cgatctcgcggtcgagctggcgcccaaggtctacgtcgctggcgcgccgcgcgcgcgcgc	687







```

QY 1273 -----cgccagagccaccgcccgcagccgctgcagggccctgctgcccgcagccgacccgc 1326
Db 1029 GAAGACCCGACCCCTGAACCCGGCGAGGATACGCCGAGCTGCTCGCGGACCGCAGCGC 1088
QY 1327 gtacagagacttcagctgagc---ggccgcgcgcagctgccccttcagagcctcga 1383
Db 1089 GTTCGAGGACTACGACGCCGACCCCTGGCGCCCAAGGCTTCGGCTTCGTCAGCTGAA 1148
QY 1384 ccagctcgcacatggaccacactgctggcgcccgccg 1419
Db 1149 CCAGCTCGCGATCGAGCACCTGCTCGGAGCCCGCTG 1184

RESULT 15
LOCUS A10241 1182 bp DNA PAT 25-JAN-1994
DEFINITION A.missouriensis DNA for glucose isomerase.
ACCESSION A10241
VERSION A10241.1 GI:490200
KEYWORDS glucose isomerase.
SOURCE Actinoplanes missouriensis.
ORGANISM Actinoplanes missouriensis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Micromonosporineae; Micromonosporaceae;
Actinoplanes.
REFERENCE 1 (bases 1 to 1182)
AUTHORS Luitzen,R.G.M., Quax,W.J., Schuurhuizen,P.W. and Mrabet,N.
TITLE Novel glucose isomerase enzymes and their use
JOURNAL Patent: EP 0351029-A 14 17-JAN-1990;
GIST-BROCADES N.V.; PLANT GENETIC SYSTEMS, N.V
FEATURES
     source
       1..1182
         /organism="Actinoplanes missouriensis"
         /db_xref="taxon:1866"
       1..>1182
         /codon_start=1
         /transl_table=11
         /product="glucose isomerase"
         /protein_id="CAA00884.1"
         /db_xref="GI:580657"
         /db_xref="SWISS-PROT:P12851"
         /translation="MSVQATREDKESFGLWTVMQARDAFGDATRLDVPVAVHKL
         EGATGIFHDDLVFPFGSDAETRGIFAGFKALDEGLIIPMTNLTFTPVFKD
         GTSNDRSVRRVIAIRKVLQMDLGAELGAKTLVLWGGREGAEDSAKOVSAALDRRE
         ALNLLAQYSEDRGYLFAIEPKNEPRGIDLLTAGHAIAFVQLEPELFGINPET
         GHEQMSNFTQGIQAALWHKKLFHDLNGOHGPKFDQDLVFGHDLNFAFLVDLLE
         NPGDAPAYDGRPHFDYKPSRTEDYGVWESAKANIRMYLLKRAKAPRAPEVOEA
         LAASKVAELKTLTPNGEGYAEALLADRSFEDYDADAVGAKGFGFVKLNQLAIEHLG
         AR"
     BASE COUNT 209 a 429 c 376 g 168 t
     ORIGIN

Query Match 38.7%; Score 598.4; DB 5; Length 1182;
Best Local Similarity 72.1%; Pred. No. 7.2e-46;
Matches 846; Conservative 0; Mismatches 301; Indels 27; Gaps 4;

QY 271 ccagccaccacccgagacaggttcaccttcggcctggtgacccgtggtggcagggcg 330
Db 9 CCAGGCCACACCGGAGACAAAGTTCCTTCGGTCTCTGACCCGTTGGATGGCAGGCTCG 68
QY 331 ggacccttcggcgagccacgcgctccccccttcgaccgggtcga---cgtcagcggt 387
Db 69 TGACCGGTTTCGGTGACGCCACCGGTACGGCACTCGACCCGTCGAGGCCGCTGCACAAAGCT 128
QY 388 ggcgaactggcgccctagagagtgaccttcacagcagcagcactgatcccttcggggc 447
Db 129 CGGTGAGATCGGGCGCTACGGCATCATCGTTCACAGCAGACGACCTGCTGCGCTTCGGCTC 188
QY 448 gtccgacaccgagcgcgagcgccacatcagcggttccttcgtagcgcctcagcgcgaccg 507
Db 189 GGACGCCACACCCGCGAGCGGATCATCGCGGGCTTCAGAAAGCGGCTTCAGCAGACCGG 248

```

Search completed: March 22, 2000, 23:29:30

Job time: 1477 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2000, 23:30:34 ; Search time 27.38 Seconds  
(without alignments)  
333.924 Million cell updates/sec

Title: US-09-383-318-2  
Perfect score: 2029  
Sequence: 1 MNQPTPEDRFTGLTWGV.....GMAFERLDQLAMHLLGARG 386

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1900	93.6	388	1 R02215	Sequence encoding
2	1867.5	92.0	387	1 W63617	S. griseofuscus g1
3	1867.5	92.0	389	1 R02216	Sequence encoding
4	1867	92.0	388	1 P94687	DNA-deduced sequen
5	1863.5	91.8	387	1 W63618	S. griseofuscus g1
6	1859	91.6	387	1 R02217	Sequence encoding
7	1744.5	86.0	371	1 R03028	Xylose isomerase g
8	1715.5	84.5	345	1 R02218	Sequence encoding
9	1530.5	75.4	394	1 P90423	Streptomyces spp.
10	1361.5	67.1	394	1 R02213	Sequence encoding
11	1353.5	66.7	394	1 R03029	Xylose isomerase g
12	1351.5	66.6	394	1 R02212	Complete sequence
13	1348.5	66.5	394	1 R13280	Glucose isomerase
14	1347.5	66.4	394	1 R05282	Amino acid sequenc
15	1345.5	66.3	394	1 R13282	Glucose isomerase
16	1344.5	66.3	394	1 R13279	Glucose isomerase
17	1342.5	66.2	394	1 R13278	Glucose isomerase
18	1341.5	66.1	394	1 R13281	Glucose isomerase
19	1341.5	66.1	394	1 R13283	Glucose isomerase
20	1338.5	66.0	394	1 R13285	Glucose isomerase
21	1333.5	65.7	394	1 R13284	Glucose isomerase
22	1332.5	65.7	394	1 R13287	Glucose isomerase
23	1323.5	65.2	394	1 R13286	Glucose isomerase
24	1311.5	64.6	391	1 R13163	Glucose isomerase
25	1282	63.2	395	1 R02214	Sequence encoding
26	1280	63.1	395	1 R03030	Xylose isomerase g
27	1124	55.4	387	1 R22623	Thermus aquaticus
28	862	42.5	435	1 R44236	Glucose isomerase
29	320	15.8	444	1 W26508	Hyperthermostable
30	275.5	13.6	440	1 R22720	Xylose isomerase x
31	275	13.6	439	1 R03024	Xylose isomerase g
32	272.5	13.4	440	1 R05693	BglII fragment con
33	255.5	12.6	440	1 R03023	Xylose isomerase g
34	245.5	12.1	479	1 R99230	Barley xylose isom

35 106 5.2 3398 1 R44430  
36 105 5.2 4572 1 W52845  
37 103 5.1 573 1 R04715  
38 99 4.9 572 1 R04716  
39 99 4.9 573 1 R64765  
40 96 4.7 800 1 P70420  
41 95.5 4.7 329 1 W23389  
42 95 4.7 921 1 W20333  
43 94.5 4.7 573 1 R64766  
44 94.5 4.7 878 1 W56116  
45 94 4.6 1896 1 W72095

eryA region polype  
A. mediterranei r1  
Amino acid sequenc  
Amino acid sequenc  
M. leprae 65 kDa p  
Sequence encoded b  
Shaker-like potass  
HSV-2 strain SB5 C  
M. tuberculosis 65  
Microbispora therm  
HSV-2 strain SB5 C

RESULT 1  
R02215  
ID R02215 standard; protein; 388 AA.  
AC R02215;  
DT 23-AUG-1990 (first entry)  
DE Sequence encoding Streptomyces murinus DSM 40091 (Smu) glucose  
KW isomerase (GI)  
KW Glucose isomerase; Streptomyces murinus DSM 40091.  
OS Streptomyces murinus.  
PN EP-351029-A.  
PD 17-JAN-1990.  
PF 17-JUL-1989; 201892.  
PR 04-NOV-1988; EP-402789.  
PA (KONN) Gist-Brocades NV (Plan-).  
PI Luiten RGM, Quax WJ, Schuurhuizen PW, Mrabet N;  
DR WPI; 90-016368/03.  
DR N-PSDB; Q93159.  
PT New mutant glucose isomerase enzymes -  
PT obtd. by site-directed mutagenesis of gene from Actinoplanes  
PT missouriensis, used for prodn. of high fructose corn syrups  
PS Disclosure; 56pp; English.  
CC The invention is a novel mutant glucose isomerase (GI) enzyme with  
CC improved properties thru AA substitn. The GI is pref. derived from  
CC Actinoplanes missouriensis. The substitn. is Lys for Arg.  
CC or vice versa. Figure 21 compares AA sequences of GI from different  
CC sources.  
SQ Sequence 388 AA;

Query Match	93.6%	Score 1900;	DB 1;	Length 388;
Best Local Similarity	92.8%	Pred. No. 3.4e-171;		
Matches 360;	Conservative 14;	Mismatches 12;	Indels 2;	Gaps 2;
QY 1	MNYQPTPEDRFTGLTWGVQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP	59		
Db 1	MSFQPTPEDRFTGLTWGVQGRDPFGDTRPALDPVETVQRLAELGAYGVTFHDDDLIP	60		
QY 60	FGASDTEREAHVKKRFOALDAGTGMTPMATTNLFTHPVFKAGFTANDRAVRYALRKTI	119		
Db 61	FGSSDTERESHKRFQALDAGTGMTPMATTNLFTHPVFDGFTANDRVRYALRKTI	120		
QY 120	RNIDLAELGAKYVYVANGGREGAAGKDVRAALDKMEAFDILLGEYVTSQGYDIRFAI	179		
Db 121	GNIDLAELGAKYVYVANGGREGAAGKDVRAALDKMEAFDILLGEYVTAQGYDIRFAI	180		
QY 180	EKPKNPRGDILLPTIGHALAFIERLERPELVGNPVEVGEHQMAGLNFPFHQAALWAGK	239		
Db 181	EKPKNPRGDILLPTVGHALAFIERLERPELVGNPVEVGEHQMAGLNFPFHQAALWAGK	240		
QY 240	LPHIDLNGSGIKYDODLRFAGDLRAAFWLYDLLESAGWEGPRHDFKPPRTEDIDGVW	299		
Db 241	LPHIDLNGSGIKYDODLRFAGDLRAAFWLYDLLESAGWEGPRHDFKPPRTEDIDGVW	300		
QY 300	ASAAGCMRNYLILKERAFAAFAADPEVQEAALRAALDQAEPTAADGLQALLADRTAYEDF	359		
Db 301	ASAAGCMRNYLILKERAFAAFAADPEVQEAALRAALDQAEPTAADGLQALLADRTAYEDF	360		

ALIGNMENTS

QY 360 DVD-AAARGMAFERLDQLAMHLLGARG 386  
 Db 361 DVDAARGMAFERLDQLAMHLLGARG 388

RESULT 2

W63617  
 ID W63617 standard; Protein: 387 AA.  
 AC W63617;  
 DT 21-SEP-1998 (first entry)  
 DE S. griseofuscus glucose isomerase.  
 KW glucose isomerase; heat resistance; thermostable; mutant;  
 OS Streptomyces griseofuscus.  
 FH Key Location/Qualifiers  
 FT Protein 2..387  
 FT /note= "mature protein"  
 FT Misc\_difference 143  
 FT /note= "encoded by GAC"  
 FT Misc\_difference 380  
 FT /note= "encoded by GAC"  
 PN J10084956-A.  
 PD 07-APR-1998.  
 PF 12-SEP-1996; 262370.  
 PR 12-SEP-1996; JP-262370.  
 PA (GODO ) GODO SHUSEI KK.  
 DR WPI: 98-264845/24.  
 DR N-PSDB: V32494.

PT Improving heat resistance of glucose isomerase - comprises replacing  
 part of amino acid in specific position of amino acid sequence  
 PS Claim 1; Fig 1; 7pp; Japanese.  
 CC This represents a Streptomyces glucose isomerase that can be used  
 to create a mutant with improved heat resistance which comprises  
 replacing an amino acid in a specific position of the sequence. The  
 glucose isomerase is useful for the preparation of isomerised sugar.  
 SQ Sequence 387 AA;

Query Match 92.0%; Score 1867.5; DB 1; Length 387;  
 Best Local Similarity 91.8%; Pred. No. 4e-168;  
 Matches 356; Conservative 16; Mismatches 13; Indels 3; Gaps 3;

QY 1 MNYQPTPEDRTFTGLTWVGWGRDPFGDTRPALDPVD-VORLAEIGAYGVTFHDDLLIP 59  
 Db 1 MSFQTPEDKFTFTGLTWVGWGRDPFGDTRPALDPVETVQRLAEIGAYGVTFHDDLLNP 60  
 QY 60 FGASDTERAHVKRFRQALDGMTVPMTNLFTHPVFKAGAFANDRAVRARYALRKT 119  
 Db 61 FGSSDTERESHKFRQALDGMTVPMTNLFTHPVFK-DRFTANDRDVRAVRKTI 119  
 QY 120 RNIDLAVELGAKVYVANGREGAESAAGKDVRAALDRMKEAFDILLGEVYTSOGYDIRFAI 179  
 Db 120 RNIDLAELGAKTYVANGREGAESGGKDVRAALDRMKEAFDILLGEVYTAQGYDIRFAI 179  
 QY 180 EPKPNPEGRDILLPTIGHALAFIERLERPELYGVNPEVGEHQMAGLNFPHGIAQALWAGK 239  
 Db 180 EPKPNPEGRDILLPTVGHALAFIERLERPELYGVNPEVGEHQMAGLNFPHGIAQALWAGK 239  
 QY 240 LFHIDLNGSGIKYDQDLRFAGAGDURAAFWLVDLLESAGWEGPRHFDKPPRTEDIDGVW 299  
 Db 240 LFHIDLNGSGIKYDQDLRFAGAGDURAAFWLVDLLESAGWEGPRHFDKPPRTEDIDGVW 299  
 QY 300 ASAAGCMRNLYLTKERAAAFRADPEVQEARAALDQLAEPTAADGLQALLADRTAYEDF 359  
 Db 300 ASAAGCMRNLYLTKERAAAFRADPEVQEARAALDQLAEPTAADGLQALLADRTAYEDF 359  
 QY 360 DVD-AAARGMAFERLDQLAMHLLGARG 386  
 Db 360 DVDAARGMAFERLDQLAMHLLGARG 387

RESULT 3

R02216

ID R02216 standard; protein; 389 AA.  
 AC R02216;  
 DT 23-AUG-1990 (first entry)  
 DE Sequence encoding Streptomyces violaceoniger CBS 409.73 glucose  
 isomerase (GI)  
 KW Glucose isomerase; Streptomyces violaceoniger CBS 409.73.  
 OS Streptomyces violaceoniger.  
 PN EP-351029-A.  
 PD 17-JAN-1990.  
 PF 17-JUL-1989; 201892.  
 PR 04-NOV-1988; EP-402789.  
 PA (KONN) Gist-Brocades NV (Plan-).  
 PI Luiten RGM, Quax WJ, Schuurhuizen PW, Mrabet N;  
 DR WPI: 90-016368/03.  
 PT New mutant glucose isomerase enzymes -  
 obtd. by site-directed mutagenesis of gene from Actinoplanes  
 PT missouriensis, used for prodn. of high fructose corn syrups  
 PS Disclosure; 56pp; English.  
 CC The invention is a novel mutant glucose isomerase (GI) enzyme with  
 improved properties thru AA substitn. The GI is pref. derived from  
 CC Actinoplanes missouriensis. The substitn. is Lys for Arg.  
 CC or vice versa. Figure 21 compares AA sequences of GI from different  
 CC sources.  
 SQ Sequence 389 AA;

Query Match 92.0%; Score 1867.5; DB 1; Length 389;  
 Best Local Similarity 91.0%; Pred. No. 4e-168;  
 Matches 354; Conservative 16; Mismatches 16; Indels 3; Gaps 2;

QY 1 MNYQPTPEDRTFTGLTWVGWGRDPFGDTRPALDPVD-VORLAEIGAYGVTFHDDLLIP 59  
 Db 1 MSFQTPEDKFTFTGLTWVGWGRDPFGDTRPALDPVETVQRLAEIGAYGVTFHDDLLIP 60  
 QY 60 FGASDTERAHVKRFRQALDGMTVPMTNLFTHPVFKAGAFANDRAVRARYALRKT 119  
 Db 61 FGSSDTERESHKFRQALDGMTVPMTNLFTHPVFKDGGFTANDRDVRAVRKTI 120  
 QY 120 RNIDLAVELGAKVYVANGREGAESAAGKDVRAALDRMKEAFDILLGEVYTSOGYDIRFAI 179  
 Db 121 RNIDLAELGAKTYVANGREGAESGGKDVRAALDRMKEAFDILLGEVYTAQGYDIRFAI 180  
 QY 180 EPKPNPEGRDILLPTIGHALAFIERLERPELYGVNPEVGEHQMAGLNFPHGIAQALWAGK 239  
 Db 181 EPKPNPEGRDILLPTVGHALAFIERLERPELYGVNPEVGEHQMAGLNFPHGIAQALWAGK 240  
 QY 240 LFHIDLNGSGIKYDQDLRFAGAGDURAAFWLVDLLESAGWEGPRHFDKPPRTEDIDGVW 299  
 Db 241 LFHIDLNGSGIKYDQDLRFAGAGDURAAFWLVDLLESAGWEGPRHFDKPPRTEDIDGVW 300  
 QY 300 ASAAGCMRNLYLTKERAAAFRADPEVQEARAALDQLAEPTAADGLQALLADRTAYEDF 359  
 Db 301 ASAAGCMRNLYLTKERAAAFRADPEVQEARAALDQLAEPTAADGLQALLADRTAYEDF 360  
 QY 360 DVDAARGMA--FERLDQLAMHLLGARG 386  
 Db 361 DVDAARGMA--FERLDQLAMHLLGARG 389

RESULT 4

P94687  
 ID P94687 standard; protein; 388 AA.  
 AC P94687;  
 DT 14-JUL-1990 (first entry)  
 DE DNA-duced sequence of Streptomyces rubiginosus xylose isomerase  
 KW Streptomyces rubiginosus xylose isomerase; muten; fructose;  
 OS Streptomyces rubiginosus  
 PN W08901520-A.  
 PD 23-FEB-1989.  
 PF 10-AUG-1988; U02765.  
 PR 22-MAR-1988; US-171693, US-084479.  
 PA (CETU) Cetus Corp (UYOR-).



PI Drummond RJ, Bloch W, Matthews BW, Toy PL, Nicholson HH;  
 DR WPI; 89-068875/09.  
 DR N-PSDB; N91083.  
 PT Increasing stability of proteins by specific aminoacid replacement -  
 PT where aminoacid introduced decreases configuration entropy of  
 PT unfolding of protein  
 PS Disclosure: ; 134pp; English.  
 CC New in the patent are mutans of the native protein given in p94687.  
 CC These mutans are used to convert glucose to fructose, and xylose to  
 CC xylulose, eg in mfr. of sweeteners. Compared with native xylose the  
 CC mutans have different chemical and thermal stability, kinetic constants,  
 CC specificity and/or lower optimum pH. Amino acid substitution sites are  
 CC selected by first determining, from the crystallographic structure, the  
 CC phi and psi angles, then screening for these for values within the specified  
 CC ranges a site which has a phi backbone configuration angle -40 to  
 CC -90 degrees when psi configuration angle is 0 to -60 degrees, or -40 to  
 CC 95 degrees when psi - 120 to 180 degrees, and is able to accommodate  
 CC AA without distortion of the 3-D protein structure, so that the  
 CC substitution decreases the configurational entropy of unfolding CEU; or  
 CC replacing a Gly residue having a negative phi angle with Ala, also  
 CC decreasing the CEU.  
 SQ Sequence 388 AA;

Query Match 92.0%; Score 1867; DB 1; Length 388;  
 Best Local Similarity 91.8%; Pred. No. 4.4e-168;  
 Matches 356; Conservative 14; Mismatches 16; Indels 2; Gaps 2;  
 QY 1 MNYQTPEDRFTFGLTWGQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDLLIP 59  
 DB 1 MNYQTPEDRFTFGLTWGQGRDPFGDTRPALDPVESVRLAELGAYGVTFHDDLLIP 60  
 QY 60 FGASSTERAHVKRFRQALDGMTVPMTNLTFTHPVKAGAFANDRAVRRLRKTI 119  
 DB 61 FGSSDSREEHVKRFRQALDGMTVPMTNLTFTHPVKAGAFANDRAVRRLRKTI 120  
 QY 120 RNIDLAELGAYGVYVANGREGAESGAADKVRALDRMKEAFDLLGEYTSOGYDIRFAI 179  
 DB 121 RNIDLAELGAYGVYVANGREGAESGAADKVRALDRMKEAFDLLGEYTSOGYDIRFAI 180  
 QY 180 EPKNEPRGDILLPTIGHALAFIERLERPELVGNPEVGEQMGALNFPHGIAQALWACK 239  
 DB 181 EPKNEPRGDILLPTIGHALAFIERLERPELVGNPEVGEQMGALNFPHGIAQALWACK 240  
 QY 240 LPHIDLNGSGIKYDQDLRFAGADLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGW 299  
 DB 241 LPHIDLNGSGIKYDQDLRFAGADLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGW 300  
 QY 300 ASAAGCMRNLYLTKERAAFRADPEVQEARLRAARDLQAEPTAADGLQALLADRTAYEDF 359  
 DB 301 ASAAGCMRNLYLTKERAAFRADPEVQEARLRAARDLQAEPTAADGLHPLDDRSFAFEF 360  
 QY 360 DVD-AAARGMAFERLDQALMDHLLGARG 386  
 DB 361 DVDAAARGMAFERLDQALMDHLLGARG 388

RESULT 5  
 ID W63618 standard; Protein; 387 AA.  
 AC W63618;  
 DT 21-SEP-1998 (first entry)  
 DE S. griseofuscus glucose isomerase mutant Y252F.  
 KW Glucose isomerase; heat resistance; thermostable; mutant;  
 OS Streptomyces griseofuscus.  
 FT Key Location/Qualifiers  
 FT Misc\_difference 253 /label= Y252F  
 FT /note= "mature wild-type Tyr is replaced with Phe"  
 PN J10084956-A.  
 PD 07-APR-1998.

PF 12-SEP-1996; 262370.  
 PR 12-SEP-1996; JP-262370.  
 PA (GODO ) GODO SHUSEI KK.  
 DR WPI; 98-264846/24.  
 PT Improving heat resistance of glucose isomerase - comprises replacing  
 PT part of amino acid in specific position of amino acid sequence  
 PS Claim 1; Page -; 7pp; Japanese.  
 CC This represents a mutant of Streptomyces glucose isomerase that has  
 CC improved heat resistance. The mutant was created by replacing an  
 CC amino acid in a specific position of the mature wild-type sequence.  
 CC The glucose isomerase is useful for the preparation of isomerised  
 CC sugar.  
 CC Note: this sequence is not provided in the specification; it has been  
 CC created by modifying the glucose isomerase sequence provided in Fig 1.  
 SQ Sequence 387 AA;

Query Match 91.8%; Score 1863.5; DB 1; Length 387;  
 Best Local Similarity 91.5%; Pred. No. 9.4e-168;  
 Matches 355; Conservative 17; Mismatches 13; Indels 3; Gaps 3;  
 QY 1 MNYQTPEDRFTFGLTWGQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDLLIP 59  
 DB 1 MSFQTPEDKFTFGLTWGQGRDPFGDTRPGLDPVETVQRLAELGAYGVTFHDDLLIP 60  
 QY 60 FGASSTERAHVKRFRQALDGMTVPMTNLTFTHPVKAGAFANDRAVRRLRKTI 119  
 DB 61 FGSSDSREESHVKRFRQALDGMTVPMTNLTFTHPVK-K-DRFTANDRDRVAYVRKTI 119  
 QY 120 RNIDLAELGAYGVYVANGREGAESGAADKVRALDRMKEAFDLLGEYTSOGYDIRFAI 179  
 DB 120 RNIDLAELGAYGVYVANGREGAESGAADKVRALDRMKEAFDLLGEYTAQYDIRFAI 179  
 QY 180 EPKNEPRGDILLPTIGHALAFIERLERPELVGNPEVGEQMGALNFPHGIAQALWACK 239  
 DB 180 EPKNEPRGDILLPTIGHALAFIERLERPELVGNPEVGEQMGALNFPHGIAQALWACK 239  
 QY 240 LPHIDLNGSGIKYDQDLRFAGADLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGW 299  
 DB 240 LPHIDLNGSGIKYDQDLRFAGADLRAAFWLVLDLESAGWEGPRHDFKPPRTEDIDGW 299  
 QY 300 ASAAGCMRNLYLTKERAAFRADPEVQEARLRAARDLQAEPTAADGLQALLADRTAYEDF 359  
 DB 300 ASAAGCMRNLYLTKERAAFRADPEVQEARLRAARDLQAEPTAADGLQALLADRTAYEDF 359  
 QY 360 DVD-AAARGMAFERLDQALMDHLLGARG 386  
 DB 360 DVDAAARGMAFERLDQALMDHLLGARG 387

RESULT 6  
 ID R02217 standard; protein; 387 AA.  
 AC R02217;  
 DT 23-AUG-1990 (first entry)  
 DE Sequence encoding Streptomyces violaceoruber LMG 7183 (Svr) glucose  
 DE isomerase (GI)  
 KW Glucose isomerase; Streptomyces violaceoruber.  
 OS Streptomyces violaceoruber.  
 PN EP-351029-A.  
 PD 17-JAN-1990.  
 PF 17-JUL-1989; 201892.  
 PR 04-NOV-1988; EP-402789.  
 PA (KONN) Gist-Brocades NV (Plan-).  
 PI Luiten RGM, Quax WJ, Schuurhuizen PW, Mrabet N;  
 DR WPI; 90-016368/03.  
 PT New mutant glucose isomerase enzymes -  
 PT obtd. by site-directed mutagenesis of gene from Actinoplanes  
 PT missouriensis, used for prodn. of high fructose corn syrups  
 PS Disclosure; ; 56pp; English.  
 CC The invention is a novel mutant glucose isomerase (GI) enzyme with  
 CC improved properties thru AA substitn. The GI is pref. derived from  
 CC Actinoplanes missouriensis. The substitn. is Lys for Arg.

```
CC or vice versa. Figure 21 compares AA sequences of GI from different
CC sources.
SQ Sequence 387 AA;

Query Match 91.6%; Score 1859; DB 1; Length 387;
Best Local Similarity 91.2%; Pred. No. 2.5e-167;
Matches 353; Conservative 15; Mismatches 17; Indels 2; Gaps 2;

QY 1 MNYOPTPEDRTFGLTWVGWGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 59
DQ 1 MNYOPTPEDRTFGLTWVGWGRDPFGDTRPALDPVDPAESVRLSELGAYGVTFHDDDLIP 60
QY 60 FGASDTEREAHVKKFRQALDAGTMTVPMTNLTFTHPVKAGAFANDRAVRRYALRRTI 119
DQ 61 FGSDTERESHKFRQALDAGTMTVPMTNLTFTHPVKDGAFTANDRDRYRVALRRTI 120
QY 120 RNIDLAVELGAKYVAVGREGSAGAAKDVRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179
DQ 121 RNIDLAVELGASVYVAVGREGSAGAAKDVRAALDRMKEAFDILLGEYVTSQGYDLKFAI 180
QY 180 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 239
DQ 181 EPKNEPRGDILLPTVGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 240
QY 240 LFHIDLNGSQGIKYDQDLRFAGDGLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVW 299
DQ 241 LFHIDLNGSQGIKYDQDLRFAGDGLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVW 300
QY 300 ASAAGCMRNLYILKERAAFRADPEVQBALRAARLDQAEPTADGLOALLADRTAYEDF 359
DQ 301 ASAAGCMRNLYILKDRAAFADPEVQBALRAARLDQAEPTADGLOALLADRTAYEDF 360
QY 360 DVD-AAARGMAFERLDQLAMDHLLGARG 385
DQ 361 DVDAARGMAFERLDQLAMDHLLGARG 387

RESULT 7
ID R03028 standard; protein; 371 AA.
AC R03028;
DE 11-JUL-1990 (first entry)
KW Xylose isomerase gene of Streptomyces violaceoniger.
OS Xylose isomerase; Arthobacter.
PN Streptomyces violaceoniger.
PD W09000198-A.
PF 04-JUL-1989; G00748.
PR 04-JUL-1988; GB-015902.
PI (BLOW) Blow D M.
PI Blow DM, Hartley BS, Henrick K;
DR WPI; 90-037131/05.
PT Xylose isomerase mutants -
PT having amino acid replacements to improve activity under acid pH
PT conditions and/or increase stability and/or affinity for metal ions.
PS Disclosure; Fig 7; 35pp; English.
CC Mutants of the xylose isomerase (XI) gene esp. Arthobacter strain B3728,
CC have been shown to have increased thermal stability, and/or improved
CC activity at low pH, and/or affinity for metal ions esp. Ca.
CC Useful in industrial conversion of glucose to sweeter sugars.
SQ Sequence 371 AA;

Query Match 85.08; Score 1744.5; DB 1; Length 371;
Best Local Similarity 87.2%; Pred. No. 1.5e-156;
Matches 340; Conservative 18; Mismatches 5; Indels 23; Gaps 6;

QY 1 MNYOPTPEDRTFGLTWVGWGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 59
DQ 1 MNYOPTPEDRTFGLTWVGWGRDPFGDTRPALDPVDVETVORLAELGAYGVTFHDDDLIP 60
QY 60 FGASDTEREAHVKKFRQALDAGTMTVPMTNLTFTHPVKAGAFANDRAVRRYALRRTI 119
```

```
DB 61 FGSDTERESHKFRQALDAGTMTVPMTNLTFTHPVKDG-FTANDRDRYRVALRRTI 119
QY 120 RNIDLAVELGAKYVAVGREGSAGAAKDVRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179
DQ 120 RNIDLAAELG-----AKDVRDALDRMKEAFDILLGEYVTAQGYDLRFAI 162
QY 180 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 239
DQ 163 EPKNEPRGDILLPTVGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 222
QY 240 LFHIDLNGSQGIKYDQDLRFAGDGLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVW 299
DQ 223 LFHIDLNGSQGIKYDQDLRFAGDGLRAAFWLVLDLLESAGY-GPRHFDKPPRTEDIDGVW 281
QY 300 ASAAGCMRNLYILKERAAFRADPEVQBALRAARLDQAEPTADGLOALLADRTAYED 358
DQ 282 ASAAGCMRNLYILKERAAFRANPEVQBALRAARLDQAEPTADGLOALLADRTAFED 341
QY 359 FDVDAARGMA--FERLDQLAMDHLLGARG 386
DQ 342 FDVEAARAARAWFERLDQLAMDHLLGARG 371

RESULT 8
ID R02218 standard; protein; 345 AA.
AC R02218;
DE 23-AUG-1990 (first entry)
KW Sequence encoding Streptomyces thermovulgaris DSM 40444 (St) glucose
DE isomerase (GI)
KW Glucose isomerase; Streptomyces thermovulgaris DSM 40444.
OS Streptomyces thermovulgaris.
PN EP-351029-A.
PD 17-JAN-1990.
PF 17-JUL-1989; 201892.
PR 04-NOV-1988; EP-402789.
PA (KONN) Gist-Brocades NV (Plan-).
PI Luiten RGM, Quax WJ, Schuurhuizen PW, Mrabet N;
DR WPI; 90-016368/03.
PT New mutant glucose isomerase enzymes -
PT obtd. by site-directed mutagenesis of gene from Actinoplanes
PT missouriensis, used for prodn. of high fructose corn syrups
PS Disclosure; 56pp; English.
CC The invention is a novel mutant glucose isomerase (GI) enzyme with
CC improved properties thru AA substitn. The GI is pref. derived from
CC Actinoplanes missouriensis. The substitn. is Lys for Arg.
CC or vice versa. Figure 21 compares AA sequences of GI from different
CC sources.
SQ Sequence 345 AA;

Query Match 84.5%; Score 1715.5; DB 1; Length 345;
Best Local Similarity 93.6%; Pred. No. 7.1e-154;
Matches 323; Conservative 14; Mismatches 7; Indels 1; Gaps 1;

QY 1 MNYOPTPEDRTFGLTWVGWGRDPFGDTRPALDPV-DVORLAELGAYGVTFHDDDLIP 59
DQ 1 MSQOPTEDRFSGLTWVGWGRDPFGDTRPALDPVGTVORLAELGAYGVTFHDDDLIP 60
QY 60 FGASDTEREAHVKKFRQALDAGTMTVPMTNLTFTHPVKAGAFANDRAVRRYALRRTI 119
DQ 61 FGAEDEREAHVKKFRQALDAGTMTVPMTNLTFTHPVKDGAFTANDRDRYRVALRRTI 120
QY 120 RNIDLAVELGAKYVAVGREGSAGAAKDVRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179
DQ 121 RNIDLAVELGARTYVAVGREGSAGAAKDVRAALDRMKEAFDILLGEYVTSQGYDIRFAI 180
QY 180 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 239
DQ 181 EPKNEPRGDILLPTVGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 240
QY 240 LFHIDLNGSQGIKYDQDLRFAGDGLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVW 299
```

```

|||||
Db 241 LFHIDLNGSGIYDQDLRFAGDLRAAFWLVLDLSSGYDGRHFDKPRPTEDLDGVW 300
QY 300 ASAAGCMRNLYLKERAAAFRAADPEVQEAALRAARLDQLAEPATAAD 344
|||||
Db 301 ASAAGCMRNLYLKERAAAFRAADPEVQEAALRAARLDQLAEPATAAD 345

RESULT 9
P90423
ID P90423 standard; protein; 394 AA.
AC P90423;
DT 1-NOV-1989 (first entry)
DE Streptomyces spp. glucose isomerase
KW Glucose isomerase; Streptomyces; DNA.
OS Streptomyces
PN J01137979-A.
PD 24-MAY-1989.
PF 24-NOV-1987; 295739.
PR 24-NOV-1987; JP-295739.
PA (NORO) Norinsho.
DR WPI; 89-198224/27.
DR N-PSDB; N90362.
PT Novel glucose isomerase gene
PT - exists in chromosome of Streptomyces and is
PT contained in recombinant etc.
PS Claim 1; fig 1; 6pp; Japanese.
CC Glucose isomerase from Streptomyces spp. (see N90362). Used to
CC study this protein's structure by recombinant techniques, and for
CC prodn. of an enzyme with improved properties.
SQ Sequence 394 AA;

Query Match 75.4%; Score 1530.5; DB 1; Length 394;
Best Local Similarity 79.0%; Pred. No. 2.3e-136;
Matches 305; Conservative 21; Mismatches 25; Indels 35; Gaps 8;

QY 1 MNYQPTPEDRFTGLTWGVQGRDPFGDATRPALDPVD-VQRLAELGAYGVTHDDDLIP 59
|||||
Db 1 MSQPTPEDKFTGLTWGVQGRDPFGDATRPGLDPVETVQRLAELGAYGVTHDDDLNP 60
QY 60 FGASDTEREAHVKKRQALDGMTVPMATNLTFTHPVKAGAFANDRAVRYALRKTI 119
|||||
Db 61 FGSDTEREASHIRKQALDGMTVPMATNLTFTHPVKAF-DRTANDRDVAYAVRKTI 119
QY 120 RNIDLAVELGAKYVAVWGREGAEGAAKDVRAALDRMKEAFDLLGEYVTSQGYDIRFAI 179
|||||
Db 120 RNIDLAELGAKYVAVWGREGAEGAAKDVRAALDRMKEAFDLLGEYVTAQGYDLRF 179
QY 180 EKPNEPRGDIILLPTTGHALAFIERLERPELYGVNPEVGEHQMAGLNFPHGIAQALWAGK 239
|||||
Db 180 EKPNEPRGDIILLPTTGHALAFIERLERPELYGVNPEVGEHQMAGLNFPHGIAQALWAGK 239
QY 240 LFHIDLNGSGIYDQDL---RGAGDLRAAFWLVLDLSSAGYEGPRHFDKPRPTEDID 296
|||||
Db 240 LFHIDLNGSGIYDQDGSRRRPAAGV----LVLDLSSAGYEGPRHFDKPRPTEDFD 295
QY 297 GWASAAGCMRNLYLKERAAAFRAADPEVQEAALRAARLDQLAEPATAAD 343
|||||
Db 296 GWASAAGCMRNLYLKERAAAFRAADPEVQEAALRAARLDQLAEPATAAD 343
QY 344 DGLQALLADRTAYEDFDVD-AAARGM 368
|||||
Db 344 DGLQALLADRTAFEDFVEAAARGM 369

```

## RESULT 10

```

R02213
ID R02213 standard; protein; 394 AA.
AC R02213;
DT 23-AUG-1990 (first entry)
DE Sequence encoding Ampullariella species ATCC 31351 glucose
DE isomerase (GI)

```

```

KW Glucose isomerase; Ampullariella species ATCC 31351.
OS Ampullariella.
PN EP-351029-A.
PD 17-JAN-1990.
PF 17-JUL-1989; 201892.
PR 04-NOV-1988; EP-402789.
PA (KONN) Gist-Brocades NV (Plan-).
PI Luiten RGM, Quax WJ, Schuurhuizen PW, Mrabet N;
DR WPI; 90-016368/03.
PT New mutant glucose isomerase enzymes -
PT obtd. by site-directed mutagenesis of gene from Actinoplanes
PT missouriensis, used for prodn. of high fructose corn syrups
PS Disclosure; 56pp; English.
CC The invention is a novel mutant glucose isomerase (GI) enzyme with
CC improved properties thru AA substitn. The GI is pref. derived from
CC Actinoplanes missouriensis. The substitn. is Lys for Arg.
CC or vice versa. Figure 21 compares AA sequences of GI from different
CC sources.
SQ Sequence 394 AA;

Query Match 67.1%; Score 1361.5; DB 1; Length 394;
Best Local Similarity 66.5%; Pred. No. 1.9e-120;
Matches 262; Conservative 45; Mismatches 78; Indels 9; Gaps 4;

QY 1 MNYQPTPEDRFTGLTWGVQGRDPFGDATRPALDPVD-VQRLAELGAYGVTHDDDLIP 59
|||||
Db 1 MSQATPDKFSFGLTWGVQGRDAFGDATRPVLDPEAVHKLAEIGAYGVTHDDDLVP 60
QY 60 FGASDTEREAHVKKRQALDGMTVPMATNLTFTHPVKAGAFANDRAVRYALRKTI 119
|||||
Db 61 FGADAATRDGIAGVFSKALDETGLIVPMVTNLTFTHPVFDGFTSNDRSRYAIRKVL 120
QY 120 RNIDLAVELGAKYVAVWGREGAEGAAKDVRAALDRMKEAFDLLGEYVTSQGYDIRFAI 179
|||||
Db 121 RQMDLGAELGAKTVLVWGREGAEGYDSAKDVGAALDRYREALNLLAQYSEDQGYGLRF 180
QY 180 EKPNEPRGDIILLPTTGHALAFIERLERPELYGVNPEVGEHQMAGLNFPHGIAQALWAGK 239
|||||
Db 181 EKPNEPRGDIILLPTTGHALAFIERLERPELYGVNPEVGEHQMAGLNFPHGIAQALWAGK 240
QY 240 LFHIDLNGSGIYDQDLRFAGDLRAAFWLVLDLSSA-----GWEGPRHFDKPRPTED 294
|||||
Db 241 LFHIDLNGSGIYDQDLRFAGDLRAAFWLVLDLSSA-----GWEGPRHFDKPRPTED 300
QY 295 IDGVWASAAGCMRNLYLKERAAAFRAADPEVQEAALRAARLDQLAEPATAADG--LQALLAD 352
|||||
Db 301 FDGVWESAKDNIRMYLLLLKERAKAFRAADPEVQEAALRAARLDQLAEPATAAD 360
QY 353 RTAYEDFDVDA-AAARGMAFERLDQLAHDHLLGAR 385
|||||
Db 361 RSFADYDADAVGAKGYGVFKLNQLAIDLHLLGAR 394

```

## RESULT 11

```

R03029
ID R03029 standard; protein; 394 AA.
AC R03029;
DT 11-JUL-1990 (first entry)
DE Xylose isomerase gene of Ampullariella sp.
KW Xylose isomerase; Arthobacter.
OS Ampullariella sp. strain 3876.
PN W09000196-A.
PD 11-JAN-1990.
PF 04-JUL-1989; G00748.
PR 04-JUL-1988; GB-015902.
PI (BLOW) Blow D M.
DR WPI; 90-037131/05.
PT Xylose isomerase mutants -
PT having amino acid replacements to improve activity under acid pH
PT conditions and/or increase stability and/or affinity for metal ions.
PS Disclosure; Fig 7; 35pp; English.

```

CC Mutants of the xylose isomerase (XI) gene esp. *Arthobacter* strain B3728,  
 CC have been shown to have increased thermal stability, and/or improved  
 CC activity at low pH, and/or affinity for metal ions esp. Ca.  
 CC Useful in industrial conversion of glucose to sweeter sugars.  
 SQ Sequence 394 AA;

```

Query Match      66.7%; Score 1353.5; DB 1; Length 394;
Best Local Similarity 66.0%; Pred. No. 1.1e-119;
Matches 260; Conservative 46; Mismatches 79; Indels 9; Gaps 4;

Qy 1 MNYQTPEDRFTGLWTVGMQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
Db 1 MSLQATPDQDKFSFGLWTVGMQARDAFGDTALDPVEAVHKLAEIGAYGITFHHDDLVP 60
Qy 60 FGASDTEREAHVKRFQALDGMTVPMTTNLFTHPVFKAGAFANDRAVRYALRKTI 119
Db 61 FGDAATRDGIVAGFSKALDETGLIVPMVTTLNLFTHPVFKDGGFTSNDRSVRYAIRKVL 120
Qy 120 RNIDLAVELGAKVYVAMGREGAESAADKDVRAALDRMKEAFDILGGEYVTSQGYDIRFAI 179
Db 121 RQMDLGAELGAKTLLVMGREGAEYDSAKDVGAALDRYREALNLLAQYSEDGGLPFAI 180
Qy 180 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNPHGIAQALWAGK 239
Db 181 EPKNEPRGDILLPTAGHAIATVQELERPELFGINPETGHEQMSNLNFTQGIQALWHKK 240
Qy 240 LPHIDLNGSGIKYDQDLRFAGGDLRAAFWLVLDLLESA-----GWEGPRHFDKPPRTED 294
Db 241 LPHIDLNGHGHKPKDQDLVFGHGDLNLFSLVDLLENGPDGAPAYDGPKHFDYKPSRTED 300
Qy 295 IDGVWASAAGCMRNLYILKERAARADPEVQEAALRAARLDQLAEPTAAG--LQALLAD 352
Db 301 FDGVWESAKDNIRMYLLKRAKAFRADPEVQEAALAEKSVDELRTPTLNPGETYADLLAD 360
Qy 353 RTAYEDFDVDA-AARGMAFERLDQLAMDHLLGAR 385
Db 361 RSAFEDYDADAVGAKGFGVFKLNQLAIDHLLGAR 394

```

```

RESULT 12
R02212
ID R02212 standard; protein; 394 AA.
AC R02212;
DE 23-AUG-1990 (first entry)
DE Complete sequence of wildtype (WT) Actinoplanes missouriensis
DE Glucose isomerase (GI)
KW Site-directed mutagenesis; glucose isomerase; Actinoplanes missouriensis;
KW high fructose corn syrup.
OS Actinoplanes missouriensis.
PN EP-351029-A.
PD 17-JAN-1990.
PF 17-JUL-1989; 201892.
PR 04-NOV-1988; EP-402789.
PA (KONN) Gist-Brocades NV (Plan-).
PI Luiten RGM, Quax WJ, Schuurhuizen PW, Mrabet N;
DR WPI: 90-016368/03.
DR N-PSDB: Q93158.
DR New mutant glucose isomerase enzymes -
PT obtd. by site-directed mutagenesis of gene from Actinoplanes
PT missouriensis, used for prodn. of high fructose corn syrups
PS Disclosure: ; 56pp; English.
CC The invention is a novel mutant glucose isomerase (GI) enzyme with
CC improved properties thru AA substitn. The GI is pref. derived from
CC Actinoplanes missouriensis. The substitn. is Lys for Arg.
CC or vice versa. Specific Lys or Arg residues are identified and site-
CC directed mutagenesis of the DNA sequence encoding the GI is performed.
CC The novel mutant gi shows 65 per cent or more sequence homology with the
CC AA sequence of WT A. missouriensis GI. Compared to wt enzyme it shows
CC higher conversion performance and improved thermostability and pH
CC stability. 394 AA;
SQ Sequence

```

```

Query Match      66.6%; Score 1351.5; DB 1; Length 394;
Best Local Similarity 65.5%; Pred. No. 1.7e-119;
Matches 258; Conservative 51; Mismatches 76; Indels 9; Gaps 4;

Qy 1 MNYQTPEDRFTGLWTVGMQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
Db 1 MSVQATREDKFSFGLWTVGMQARDAFGDTALDPVEAVHKLAEIGAYGITFHHDDLVP 60
Qy 60 FGASDTEREAHVKRFQALDGMTVPMTTNLFTHPVFKAGAFANDRAVRYALRKTI 119
Db 61 FGSAQTRDGIAGFKALDETGLIVPMVTTLNLFTHPVFKDGGFTSNDRSVRYAIRKVL 120
Qy 120 RNIDLAVELGAKVYVAMGREGAESAADKDVRAALDRMKEAFDILGGEYVTSQGYDIRFAI 179
Db 121 RQMDLGAELGAKTLLVMGREGAEYDSAKDVSAALDRYREALNLLAQYSEDGGLPFAI 180
Qy 180 EPKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNPHGIAQALWAGK 239
Db 181 EPKNEPRGDILLPTAGHAIATVQELERPELFGINPETGHEQMSNLNFTQGIQALWHKK 240
Qy 240 LPHIDLNGSGIKYDQDLRFAGGDLRAAFWLVLDLLESA-----GWEGPRHFDKPPRTED 294
Db 241 LPHIDLNGHGHKPKDQDLVFGHGDLNLFSLVDLLENGPDGAPAYDGPKHFDYKPSRTED 300
Qy 295 IDGVWASAAGCMRNLYILKERAARADPEVQEAALRAARLDQLAEPT--AADGLQALLAD 352
Db 301 YDGVWESAKANIRMYLLKRAKAFRADPEVQEAALAEKSVDELRTPTLNPGEVYALLAD 360
Qy 353 RTAYEDFDVDA-AARGMAFERLDQLAMDHLLGAR 385
Db 361 RSAFEDYDADAVGAKGFGVFKLNQLAIEHLLGAR 394

```

```

RESULT 13
R13280
ID R13280 standard; Protein; 394 AA.
AC R13280;
DE 16-OCT-1991 (first entry)
DE Glucose isomerase mutant A25K.
KW GI, high fructose corn syrup; sweetener; sugar.
OS Actinoplanes missouriensis.
FH Key Location/Qualifiers
FT domain 10..14
FT /label= beta_strand
FT domain 49..52
FT /label= beta_strand
FT domain 87..91
FT /label= beta_strand
FT domain 132..137
FT /label= beta_strand
FT domain 177..181
FT /label= beta_strand
FT domain 211..218
FT /label= beta_strand
FT domain 241..246
FT /label= beta_strand
FT domain 289..292
FT /label= beta_strand
FT domain 35..47
FT /label= helix
FT domain 64..83
FT /label= helix
FT domain 108..130
FT /label= helix
FT domain 150..173
FT /label= helix
FT domain 195..204
FT /label= helix
FT domain 127..239
FT /label= helix
FT domain 264..276
FT /label= helix

```

```

FT domain 300..327
FT FT /label= helix
FT FT 25
FT FT /label= mutation
FT FT /note=" Ala -> Lys "
PN EP-440273-A.
PD 07-AUG-1991.
PF 02-JAN-1991; 200003.
PR 04-JAN-1990; EP-200029.
PR 02-JAN-1991; EP-200003.
PA (KONN ) GIST-BROCADES NV.
PA (PLAN-) PLANT GENETIC SIST NV.
PI Lambelir AVR, Quax WJ, Lasters I, Van der Laan JM;
DR WPI: 91-231993/32.
PT Glucose isomerase mutants with altered substrate specificity -
PT and methods for selecting aminoacid(s) for substitution, used
PT e.g. in prodn. of high fructose corn syrup.
PS Claim 13; Page 13; 18pp; English.
CC The analogue was prepd. by site directed mutagenesis of the GI
CC gene. The substn. of Lys for Ala at posn 25 introduces a positive
CC charge at a distance of 6-8A of the 01, 02 and 03 positions of the
CC substrate in the enzyme-sorbitol-cobalt complex. This causes
CC disruption of the water structure in the interface and displaces
CC 26phe, which shapes the hydrophobic pocket accommodating the C1
CC aliphatic hydrogens of the substrate. This gives the enzyme a
CC better Km(xylose)/Km(glucose) ratio compared with the wild type.
CC The GI is used in industrial processes, e.g. for the prodn. of
CC high fructose corn syrup.
CC See also R13278-R13287.
SQ Sequence 394 AA;

Query Match 66.5%; Score 1348.5; DB 1; Length 394;
Best Local Similarity 65.5%; Pred. No. 3.2e-119;
Matches 258; Conservative 50; Mismatches 77; Indels 9; Gaps 4;

QY 1 MNYQTPEDRFTFGLTWGVQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
DB 1 MSVQATREDKFSFGLTWGVQGRDPFGDTRPALDPVEAVHKLAEIGAYGVTFHDDDLIP 60
QY 60 FGASDERAHVRRFRQALDGMTVPMATTNLFTHPVFKAGAFANDRAVRYALRKTI 119
DB 61 FGSDAQTRDGIAGFKKALDEGLIVPMVTNLFTHPVFKAGAFANDRAVRYALRKVI 120
QY 120 RNIDLAVELGAKVYVANGREGAESGAADVRAALDRKAEFDLLGEYVTSOGYDIRFAI 179
DB 121 RQMDLGAELGAKTLVWGREGAEYDSKDVSAALDRYREALNLLAQYSEDGRGYGLRF 180
QY 180 EPKPNPRGDIILLPTIGHALAFIERLERPELVGNPEVGEOMAGLNFPHGIAQALW 239
DB 181 EPKPNPRGDIILLPTAGHAIQVQLERPELVGNPEVGEOMAGLNFPHGIAQALW 240
QY 240 LFHIDLNGSGIKYDQDLRFAGDRLAAFWLVDLLESA-----GWEGPRHFDK 294
DB 241 LFHIDLNGSGIKYDQDLRFAGDRLAAFWLVDLLESA-----GWEGPRHFDK 300
QY 295 IDGVWASAGCMRNLYLILKERAAFRADPEVQEAALRAALDQALAEPT--AADGL 352
DB 301 YDGVWESAKANIRMYLLKRAKAFRADPEVQEAALRAALDQALAEPT--AADGL 360
QY 353 RTAYEDFDVDA-AARGMAFERLDQALMDHLLGAR 385
DB 361 RSAFEYDADAVGAKGFGFKLNQLNLAIEHLLGAR 394

RESULT 14
R05282
ID R05282 standard; protein; 394 AA.
AC R05282;
DT 20-AUG-1990 (first entry)
DE Amino acid sequence of Actinoplanes missouriensis (DSN 43046)
DE D-glucose isomerase (EcoAmi(DSM) GI)
KW Actinoplanes missouriensis (DSM 43046); D-glucose isomerase;

KW D-xylose isomerase; D-xylose.ketol-isomerase.
OS Actinoplanes missouriensis.
PN EP355039-A.
PD 21-FEB-1990.
PF 17-JUL-1989; 201893.
PR 17-JUL-1989; EP-201893, EP-201539.
PA (KONN) Gist-Brocades NV (PLAN-).
PI Mrabet N, Lasters I, Stanssens P, Mathysens G, Wodak S, Quax WJ;
DR WPI: 90-052789/08.
DR N-PSDB; Q01862.
PT Biologically active modified protein prodn. -
PT comprises substituting arginine residue for lysine in starting protein
PS Disclosure; Fig 15; 64pp; English.
CC It is D-glucose isomerase from Actinoplanes missouriensis
CC (DSM 43046). Its DNA is genetically engineered to code for a mutein
CC that has increased thermostability and increased resistance
CC towards chemical modification. Its DNA is used in an example of
CC the method, which is claimed, where one or more Lys are substnd. by Arg or
CC vice versa, at a site which can accommodate such a substitution without
CC altering the biological activity.
SQ Sequence 394 AA;

Query Match 66.4%; Score 1347.5; DB 1; Length 394;
Best Local Similarity 65.2%; Pred. No. 4e-119;
Matches 257; Conservative 52; Mismatches 76; Indels 9; Gaps 4;

QY 1 MNYQTPEDRFTFGLTWGVQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
DB 1 VSVQATREDKFSFGLTWGVQGRDPFGDTRPALDPVEAVHKLAEIGAYGVTFHDDDLIP 60
QY 60 FGASDERAHVRRFRQALDGMTVPMATTNLFTHPVFKAGAFANDRAVRYALRKTI 119
DB 61 FGSDAQTRDGIAGFKKALDEGLIVPMVTNLFTHPVFKAGAFANDRAVRYALRKVI 120
QY 120 RNIDLAVELGAKVYVANGREGAESGAADVRAALDRKAEFDLLGEYVTSOGYDIRFAI 179
DB 121 RQMDLGAELGAKTLVWGREGAEYDSKDVSAALDRYREALNLLAQYSEDGRGYGLRF 180
QY 180 EPKPNPRGDIILLPTIGHALAFIERLERPELVGNPEVGEOMAGLNFPHGIAQALW 239
DB 181 EPKPNPRGDIILLPTAGHAIQVQLERPELVGNPEVGEOMAGLNFPHGIAQALW 240
QY 240 LFHIDLNGSGIKYDQDLRFAGDRLAAFWLVDLLESA-----GWEGPRHFDK 294
DB 241 LFHIDLNGSGIKYDQDLRFAGDRLAAFWLVDLLESA-----GWEGPRHFDK 300
QY 295 IDGVWASAGCMRNLYLILKERAAFRADPEVQEAALRAALDQALAEPT--AADGL 352
DB 301 YDGVWESAKANIRMYLLKRAKAFRADPEVQEAALRAALDQALAEPT--AADGL 360
QY 353 RTAYEDFDVDA-AARGMAFERLDQALMDHLLGAR 385
DB 361 RSAFEYDADAVGAKGFGFKLNQLNLAIEHLLGAR 394

RESULT 15
R13282
ID R13282 standard; protein; 394 AA.
AC R13282;
DT 16-OCT-1991 (first entry)
DE Glucose isomerase mutant T90S.
KW GI; high fructose corn syrup; sweetener; sugar.
OS Actinoplanes missouriensis.
FH Key Location/Qualifiers
FT domain 10..14
FT FT /label= beta_strand
FT FT 49..52
FT FT /label= beta_strand
FT FT 87..91
FT FT /label= beta_strand
FT FT 132..137
FT FT /label= beta_strand

```

FT domain 177..181  
 FT /label= beta\_strand  
 FT domain 211..218  
 FT /label= beta\_strand  
 FT domain 241..246  
 FT /label= beta\_strand  
 FT domain 289..292  
 FT /label= beta\_strand  
 FT domain 35..47  
 FT /label= helix  
 FT domain 64..83  
 FT /label= helix  
 FT domain 108..130  
 FT /label= helix  
 FT domain 150..173  
 FT /label= helix  
 FT domain 195..204  
 FT /label= helix  
 FT domain 127..239  
 FT /label= helix  
 FT domain 264..276  
 FT /label= helix  
 FT domain 300..327  
 FT /label= helix  
 FT region 186  
 FT /label= mutation  
 FT /note=" Glu -> Gln "  
 PN EP-440273-A.  
 PD 07-AUG-1991.  
 PF 02-JAN-1991; 200003.  
 PR 04-JAN-1990; EP-200029.  
 PR 02-JAN-1991; EP-200003.  
 PA (KONN ) GIST-BROCADES NV.  
 PA (PLAN-) PLANT GENETIC SYST NV.  
 PI Lambelir AVR, Quax WJ, Lasters I, Van der Laan JM;  
 DR WPI; 91-231993/32.  
 PT Glucose isomerase mutants with altered substrate specificity -  
 PT and methods for selecting aminoacid(s) for substitution, used  
 PT e.g. in prodn. of high fructose corn syrup.  
 PS Claim 13; Page 13; 18pp; English.  
 CC The analogue was prepd. by site directed mutagenesis of the GI  
 CC gene. The substn. of Glu at posn 186 gives the enzyme  
 CC a better Km(xylose)/Km(glucose) ratio than the wild type.  
 CC The GI is used in industrial processes, e.g. for the prodn. of  
 CC high fructose corn syrup.  
 CC See also R13278-R13287.  
 SQ Sequence 394 AA;

Query Match 66.3%; Score 1345.5; DB 1; Length 394;  
 Best Local Similarity 65.2%; Pred. No. 6.2e-119;  
 Matches 257; Conservative 51; Mismatches 77; Indels 9; Gaps 4;  
 QY 1 MNYQTPEDRFTFTGLWTVGMQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59  
 DB 1 MSVQATREDKFSGLWTVGMQARDAGDTRTALDPVEAVHKLAEIGAYGTTFHDDDLVP 60  
 QY 60 FGASDTEREAHVRRQALDAGTGMVPMATNTLFTHPVKAGFTANDRAVRYALRKTI 119  
 DB 61 FGSDAQTRDGIAGFKKALDETGLIVPMVTNLTFTHPVKGFGFTSNDNSVRYAIRKVL 120  
 QY 120 RNIDLAVELGAKVYVAMGREGAEGSAKDVRAALDRKAEFDLLGEYVTSQGYDIRFAI 179  
 DB 121 RQMDLGAELGAKTLVLMGREGAEGYDSAKDVSAAALDRYREALNLLAQYSEDGRGYGLRFAI 180  
 QY 180 EPKPNPRGDIILLPTIGHALAFIERLERPELYGVNPEVGEHOMAGLNPHGIAQALWACK 239  
 DB 181 EPKPNQPRGDIILLPTAGHAIAFVQELERPELFGINPETGHEQMSNLNETQIAQALWHKK 240  
 QY 240 LFHIDLNGSGIKYDQDLRFAGDGLRAAPMLVDLLESA----GWEGPRHFDKPPRTED 294  
 DB 241 LFHIDLNGQHGPKFDODLVFGHGDLLNLFSLVDLLENGPDGAPAYDGPGRHFDKPSRTED 300

QY 295 IDGVWASAGCMRNLYLILKERAAAFRADPEVQVQALRAARLDQLAEPT--AADGLQALLAD 352  
 DB 301 YDGVWESAKANIRMYLLKRAKAFRADPEVQVQALAAASKVAELKTPTLNPGGYAELLAD 360  
 QY 353 RTAYEDEFDVA-AARGMAFERLDQLAMDHLLGAR 385  
 DB 361 RSFEDYDADAVGAKGFGFVKLNQLAIEHLLGAR 394  
 Search completed: March 22, 2000, 23:43:24  
 Job time: 770 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2000, 23:43:59 ; Search time 39.33 Seconds  
(without alignments)  
680.472 Million cell updates/sec

Title: US-09-383-318-2  
Perfect score: 2029  
Sequence: 1 MNYQPTDRFTFGLTWGV.....GMAFERLDQLAMHLLGARG 386

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 59334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0  
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SPTREMBL12.\*  
1: sp.archaea.\*  
2: sp.bacteria.\*  
3: sp.fungi.\*  
4: sp.human.\*  
5: sp.invertebrate.\*  
6: sp.mammal.\*  
7: sp.mhc.\*  
8: sp.organelle.\*  
9: sp.phage.\*  
10: sp.plant.\*  
11: sp.rodent.\*  
12: sp.virus.\*  
13: sp.vertebrate.\*  
14: sp.unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2029	100.0	386	2 Q9ZAI3	Q9ZAI3 streptomyc
2	317	15.6	444	2 Q9X125	Q9X125 thermotoga
3	307	15.1	439	2 P77995	P77995 thermotoga
4	303	14.9	439	2 Q9X422	Q9X422 lactococcus
5	300.5	14.8	439	2 Q9X416	Q9X416 lactococcus
6	176	8.7	169	10 Q9XFD3	Q9XFD3 oryza sativ
7	138.5	6.8	383	2 Q9X0G0	Q9X0G0 thermotoga
8	131.5	6.5	386	2 Q9XAB3	Q9XAB3 streptomyc
9	125.5	6.2	315	2 Q69950	Q69950 streptomyc
10	120	5.9	290	2 Q9Z420	Q9Z420 streptomyc
11	107	5.3	233	1 Q26349	Q26349 methanobact
12	105	5.2	397	5 Q17899	Q17899 caenorhabdi
13	105	5.2	4725	2 Q54566	Q54566 amycolatops
14	100.5	5.0	7576	2 Q9ZGA4	Q9ZGA4 streptomyc
15	97	4.8	373	2 Q9ZC00	Q9ZC00 streptomyc
16	97	4.8	501	1 Q57693	Q57693 thermoprote
17	96	4.7	415	2 Q55407	Q55407 synecocyst
18	95.5	4.7	353	4 Q43659	Q43659 homo sapien
19	95.5	4.7	367	4 Q13303	Q13303 homo sapien
20	95.5	4.7	370	2 Q9XDNO	Q9XDNO salmonella

21 95 4.7 270 2 Q9WYP7 Q9WYP7 thermotoga  
22 95 4.7 293 2 Q66289 Q66289 agrobacteri  
23 94.5 4.7 878 2 Q9WXH8 Q9WXH8 microbisp  
24 94 4.6 287 2 P73599 P73599 synecocyst  
25 94 4.6 541 2 O50491 O50491 streptomyc  
26 94 4.6 3122 12 P89459 P89459 herpes simp  
27 94 4.6 3590 2 Q43365 Q43365 bordetella  
28 93.5 4.6 480 2 O05438 O05438 mycobacteri  
29 93.5 4.6 768 2 O086766 O086766 streptomyc  
30 93.5 4.6 1361 2 O86617 O86617 streptomyc  
31 93 4.6 318 2 Q9WX20 Q9WX20 streptomyc  
32 93 4.6 821 2 O59241 O59241 bacillus sp  
33 92.5 4.6 367 6 Q27955 Q27955 bos taurus  
34 92.5 4.6 367 11 P97381 P97381 mus musculu  
35 92.5 4.6 890 2 Q53582 Q53582 streptomyc  
36 92.5 4.6 1279 2 O50314 O50314 chlorobium  
37 91.5 4.5 678 2 Q47801 Q47801 enterococc  
38 91.5 4.5 4472 2 O33954 O33954 streptomyc  
39 91 4.5 328 1 Q9YBV2 Q9YBV2 aeropyrum p  
40 91 4.5 336 2 O06334 O06334 mycobacteri  
41 91 4.5 930 1 Q9YBL5 Q9YBL5 aeropyrum p  
42 91 4.5 3729 2 O33956 O33956 streptomyc  
43 90.5 4.5 367 11 Q64284 Q64284 mus musculu  
44 90.5 4.5 451 2 Q9X8D0 Q9X8D0 streptomyc  
45 90.5 4.5 798 2 O87626 O87626 neisseria f

## ALIGNMENTS

RESULT 1  
Q9ZAI3 PRELIMINARY: PRT: 386 AA.  
ID Q9ZAI3;  
AC Q9ZAI3;  
DT 01-MAY-1999 (Tremblrel. 10, (Created)  
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
DE XYLOSE ISOMERASE (EC 5.3.1.5).  
GN XYLA.  
OS Streptomyces sp.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SK;  
RA BELGITH-SRIH K.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: D-XYLOSE -> D-XYLULOSE.  
CC -!- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY OF THE PROTEIN.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
DR EMBL; Y15518; CAA75672.1; -.  
DR HSSP; P15587; 1XYC.  
DR PROSITE; PS00172; XYLOSE-ISOMERASE\_1; 1.  
DR PROSITE; PS00173; XYLOSE-ISOMERASE\_2; 1.  
KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium.  
SQ SEQUENCE 386 AA; 42702 MW; 7837ACF1 CRC32;

Query Match 100.0%; Score 2029; DB 2; Length 386;  
Best Local Similarity 100.0%; Pred No. 2.5e-143;  
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNYQPTDRFTFGLTWGVQGRDPFGDTRPALDPVDVQRLAELGAYGVTFHDDLLIPF 60  
|||||  
Db 1 MNYQPTDRFTFGLTWGVQGRDPFGDTRPALDPVDVQRLAELGAYGVTFHDDLLIPF 60  
|||||  
QY 61 GASDTRERAHVKFRQALDAGTMTVPMTNLTTHPVFKAGATANDRAVRRALRKRTIR 120  
|||||  
Db 61 GASDTRERAHVKFRQALDAGTMTVPMTNLTTHPVFKAGATANDRAVRRALRKRTIR 120  
|||||  
QY 121 NIDLAVELGAKVYVANGREGESGAAGKOVRAALDRMKEAFDILLGEVTSQGDYIRFAIE 180  
|||||  
Db 121 NIDLAVELGAKVYVANGREGESGAAGKOVRAALDRMKEAFDILLGEVTSQGDYIRFAIE 180  
|||||





Db	21	WAYNSGSTRFKVFAQGVPRDPF-----EKLD--DAAKVHEFTGAAPTVALH-----IPMD	69
Qy	62	ASD--TEREAHVRFRRQALDATGMTVPMTNLTHPVFKAGAFANDRAVRRYALRKTI	119
Db	70	RVEDYAALAAHAEK-----RGVRIGAINSTFQDDYRLGSGICHDPDAAVRRKAVDHL	122
Qy	120	RNIDLAVELGAKVVFVANGREGAESGAACKDVRAALDRMKEAFDLLGEYVTSQGYDIRFAI	179
Db	123	ECVDIMDATSGRLKWFPA-DGTYNPGQDDIRSRQDRLAG---LAEVYERLGEQGMILL	178
Qy	180	EPKENEP---RGDILLPTIGHALAFTERLERPELYGVNPEVGHQEMAGLNFPHGIAQALW	236

```
Db 179 EYLFEPFAFTTDV--PDWGTATAC--LKLGEKAQVVDVDTGH-HAPGNTNIEFVATLLR 233
QY 237 AGKLFHIDLNGSGIKYDQDLRFAGDLRAAFWLIV-DLLESAGWEGPRHFDFKPPRTEDI 295
Db 234 EGKLGGFDFN--SRFYADDDLVGAADPFOLFPRIMYEVVGGGF-----TSDV 279
QY 296 DGWVASAGCMRNVLILKERAAAFRADPEVOEALRAARL---DQLAEPTAA-DGLQALLA 351
Db 280 -----AFMLDCHNIEAKIPAIRSYMNVQEATKALLVDGTLALAEQAAGDVLENAV 333
QY 352 DRTAY 356
Db 334 LMDAY 338

RESULT 9
O69950 PRELIMINARY; PRT; 315 AA.
AC O69950;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE HYPOTHETICAL 33.5 KD PROTEIN.
GN SC3P9.10.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA SEGER K.J., HARRIS D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE; 97000351.
RA REDENBACH M., KIESER H.M., DENAPAITE D., EICHNER A., CULLUM J.,
RA KINASHI H., HOPWOOD D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL023862; CAAL9633.1; -.
KW Hypothetical protein.
SQ SEQUENCE 315 AA; 33543 MW; AB456D17 CRC32;

Query Match 6.2%; Score 125.5; DB 2; Length 315;
Best Local Similarity 24.4%; Pred. No. 0.062;
Matches 86; Conservative 46; Mismatches 145; Indels 75; Gaps 18;

QY 21 QGRDPFGDTRPALDPV-----DVRDLAELG--AYGVTFHDDDLIPFGA 62
Db 7 QQNGAGDG--PATHPLFCYGTNGGLADRLDDALLADLGDYGVGLTLDHMLDPLA- 64
QY 63 SDTEREAHVFRQALDATGMTVPMAT-TNLFTHPVFKAGAFAN----DRAVRRVALRK 117
Db 64 --DDLAARTRLARRLDTLGLVTGTGARYVLDPRKHGSLDPPQDRA--RTGL-- 119
QY 118 TIRNIDLAVELGAKVYVANGREGSAGKAADVRAALDRMKEAFDILLGEVTSQGYDIRF 177
Db 119 LLRAVDAAELGAHVHCFSG-----VTPGTDDEDTAKWRLAEALAPVLDAAATAG--VPL 172
QY 178 AIEPKNEPRGDTLLPTIGHALAFIERLERPELYGVNPEVGEQMAGLNPHGIAQALWA 237
Db 173 AVEPEFGH-----LLATVADFHTLGAIGDPEHLGLTLDIGHCQCCLPEPPADCVRAA-A 226
QY 238 GKLFHDLNG-QSGIKYDQDLRFAGDLRAAFWLVDLLESAGWEG-----PRHDFKPP 290
```

```
Db 227 PMLRHVOIEDMRGV--HEHLPFGDEIDFP-FVLEALATGYQGLTVVLPDRHSHAGPH 283
QY 291 RTEDIDGVWASAGCMRNVLILKERAAAFRADPEVOEALRAARLDQLAEPTA 342
Db 284 YAE-----RSLPFLRRAPAPPPP-----RTNRSGEPSA 313

RESULT 10
O92420 PRELIMINARY; PRT; 290 AA.
AC O92420;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE PUTATIVE ISOMERASE.
GN SCE29.10C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA SAUNDERS D.C., HARRIS D.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA BENTLEY S.D., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE; 97000351.
RA REDENBACH M., KIESER H.M., DENAPAITE D., EICHNER A., CULLUM J.,
RA KINASHI H., HOPWOOD D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL035707; CAB38885.1; -.
KW Isomerase.
SQ SEQUENCE 290 AA; 31122 MW; D8096EF4 CRC32;

Query Match 5.9%; Score 120; DB 2; Length 290;
Best Local Similarity 23.9%; Pred. No. 0.14;
Matches 72; Conservative 48; Mismatches 133; Indels 48; Gaps 17;

QY 22 GRDPFGDTRPALDPVDVORLAELG--AYGVTFHDDDLIPFGASDTEREAHVFRQALD 79
Db 10 GTNGFGD--HPLEDALAV--LADLGYEGVGLTLDPRHLDPF--ADDLP--RLRRLAARD 62
QY 80 ATGMTVPMATTNLT-----HPVEKAGAFANDRAVRYALRKTIKIDLAVELGAKV 132
Db 63 RUGLAVVVTGGRYVLDPRKHQPVLMSA-----EGAGRRVDL--LLRAVRIAADLGAEE 115
QY 133 YVANGREGSAGKAADVRAALDRMKEAFDILLGEVTSQGYDIRFAIEPKNEPRGDTLL 192
Db 116 VSPWGAAPADTPR---QVWDRLLAGCGTVVEAARAGVVLGF--EPEPG-----MFV 164
QY 193 PTIGHALAFIERLERPELYGVNPEVGEQMAGLNPHGIAQAL--WAGKLFHIDLNGSQS 250
Db 165 DTLDAYDELCCRLLGGPHPLGLTLDIGHRCLE---POPVADCVRRVADRLVNVOIEDMR 221
QY 251 IKYDQDLRFAGDLRAAFWLVDLLESAGWEG-----PRHDFKPP-RTEDIDGVWASAA 303
Db 222 GTEH-LEFGSGEIDFP-FVLSALAAATGYRLSVLPELPHSHAAPVARRSLDFLRAAD 279
QY 304 G 304
Db 280 G 280
```

RESULT 11  
Q26349 PRELIMINARY; PRT; 233 AA.  
AC Q26349;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)  
DE CONSERVED PROTEIN.  
GN MTH247.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
OC Methanobacterium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DELTA H;  
RX MEDLINE; 98037514.  
RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,  
RA ALDREDGE T., BASHIRADEH R., BLAKELY D., COOK R., GILBERT K.,  
RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTHIER B., QIU D.,  
RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,  
RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,  
RA MCDUGALL S., SHMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,  
RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;  
RT \*Complete genome sequence of Methanobacterium thermoautotrophicum  
RT deltaH: functional analysis and comparative genomics.\*;  
RL J. Bacteriol. 179:7135-7155(1997).  
DR EMBL; AE000811; AAB84753.1; F9D4C1C3 CRC32;  
SQ SEQUENCE 233 AA; 25067 MW; F9D4C1C3 CRC32;

Query Match 5.3%; Score 107; DB 1; Length 233;  
Best Local Similarity 26.6%; Pred. No. 0.96;  
Matches 54; Conservative 32; Mismatches 83; Indels 34; Gaps 10;  
QY 96 PVFKAGATANDRAVRRYALRKTINIDLAVELGAKYVWGGREGSAGAAKDVRAALD 155  
DB 57 PISDINIASLND-TYREASIREVASVDLAVELSDVVVHPGTVPFLGRPYRDLIA--E 113  
QY 156 RMKEAFDILLGELVTSQGYDIRFAIPKPNPR--GDILLPTIGHALAFIERLERELXGV 213  
DB 114 RNLESIAISEVASDRG-----GVVPE-NMPLEGPLL-----RELELWGV 155  
QY 214 NPEVGEQMGALNFFPHGI-----AQAALWAGKLFHIDLNGSQGIKYDQD-LRFGAGDLRAA 267  
DB 156 AEELG--VNVTLDAHAATMGVSTEEVSAHVGHVLSLSDTEVDSDHDLGSGSLDFQA- 213  
QY 268 FWLVLLLESAGWEGPRHDFKPP 290  
DB 213 --LLDGLRGAGYTGVLTVETKTP 233

RESULT 12  
Q17899 PRELIMINARY; PRT; 397 AA.  
AC Q17899;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)  
DE C10C5.4 PROTEIN.  
GN C10C5.4.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MATTHEWS P.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z68214; CAA92447.1; -;  
DR PFAM; PF01546; Peptidase\_M20; 1.  
SQ SEQUENCE 397 AA; 45242 MW; 0963D445 CRC32;

Query Match 5.2%; Score 105; DB 5; Length 397;

Best Local Similarity 20.5%; Pred. No. 2.9;  
Matches 68; Conservative 46; Mismatches 115; Indels 102; Gaps 18;  
QY 83 MTVP-----MATTNLFTH----PVFKA-----GATANDRAVRRYALRKT-----IRNI 122  
DB 58 MTIPGSELPESIMLYSHTDVVFTRHWHDPYSAFKDEGNIFARGAQDTKCLGVQHV 117  
QY 123 DLAVELGAK-----VYVWANGREGAESGAADVRAALD-----RMKEAFDILLGELV 169  
DB 118 EAFNLFAKAGAKQWKRIHVWGPDE--ETGHINGMKGFVETNEFKLINIGFSLDEGHS 175  
QY 170 SQGYDIRFAIE-----PKPNEPRGDILLPTIGHALAFIE-----R 204  
DB 176 ENGVLTFYAEVTVWLVKVTAPGNP-----GHGSQFMENTAMEKIERFLASARA 226  
QY 205 LERELGVNP--EVGHEOMAGLNFPHGIAQALWAGKLF--HIDLNGSQGIKYDQDLRFG 260  
DB 227 NEQKELLEKNPTWALGDVTTLNVLNKGVGQVFNVIPEKFEAYVDIR-----LTPNODF--- 280  
QY 261 AGDLRAAF--WLVDLLESAGWEGPRHDFKPP--PRTEIDIGWAS----- 302  
DB 280 -GEIRAKLDQWVKDAGSGVTFEFSQHSYDKPKVSPHTRD-DFWAAAFEDSLKQENCEPTTE 337  
QY 302 -AAGCMRNLYLKERAAAFRADPEVQEAALRA 331  
DB 338 VCVGSTDSTRFVRKAGVPAINFSPMINTPVA 368

RESULT 13  
Q54666 PRELIMINARY; PRT; 4735 AA.  
AC Q54666;  
DT 01-JUN-1998 (TReMBLrel. 06, Created)  
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)  
DE POLYKETIDE SYNTHASE.  
GN RIFA.  
OS Amycolatopsis mediterranei.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae; Amycolatopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S699;  
RA KIM C.G., YU T.W., FRYHLE C., HANDA S., FLOSS H.G.;  
RL J. Biol. Chem. 0:0-0(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S699;  
RA AUGUST P.R., TANG L., YOON Y.J., NING S., MUELLER R., YU T.W.,  
RA TAYLOR M., HOFFMANN D., KIM C.G., ZHANG X., HUTCHINSON C.R.,  
RA FLOSS H.G.;  
RL Chem. Biol. 5:0-0(0002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S699;  
RA AUGUST P.R., TANG L., YOON Y.J., NING S., MUELLER R., HUTCHINSON C.R.,  
RA FLOSS H.G.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-LBG A3136;  
RA SCHUPP T., TOUPET C., ENGEL N., GOFF S.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF040570; AAC01710.1; -;  
DR EMBL; AJ223012; CAAL1035.1; -;  
DR HSSP; P08659; ILIC.  
DR PROSITE; PS00455; AMP\_BINDING; 1.  
DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; 3.  
DR PFAM; PF00550; pp-binding; 4.  
DR PFAM; PF00109; ketoacyl-synt; 3.  
DR PFAM; PF00698; Acyl\_transf; 3.  
DR PFAM; PF00501; AMP-binding; 1.  
DR PRINTS; PR00154; AMPBINDING.

KW Transferase. 4735 AA; 495213 MW; FC634DAE CRC32;

Query Match 5.2%; Score 105; DB 2; Length 4735;

Best Local Similarity 24.5%; Pred. No. 1,1e-02;

Matches 92; Conservative 46; Mismatches 140; Indels 98; Gaps 22;

QY 3 YQPTPEDRTFF-----GLTVGWQGRDPFGDTRPALDP-VDVQRLAEAGVGTFFHDD 56

DB 4417 FETDPLDAFVFSIAGGGGQ--PAG-RANAVLDALVEWRRARGLKATSIANGALD 4473

QY 57 LIPFGASDTERAHYKRFQALDAGTMTVPMATNLTTHPVFKAGAFANDRAVRYALR 116

DB 4474 QIGIGMD-----EAALAQLR--GVIPMAPPLAVTAM-----VQAVAGNEKAV----- 4516

QY 117 KTRINIDLAVELGAKVYVWGGREG---AESGAADV-RAALDRMKEAFDILLGEYVTSQ 172

DB 4516 -AVADMWRAFIPTSV-----RPSPLFADLPEAKAILRAAQDDGD-----GDTASSLA 4565

QY 173 YDIRFAIEPKPNRPGDILLPTI-GHALAFIERLERPELYGVNP-----EVGHEQMAGLN 226

DB 4566 DSLRAV-----PDAEQNRILLKVRGHASTV---LHSGAEGIGPRQAFQEVGFDSLAAYN 4618

QY 227 FPHGTAQALWACKLPHIDLNGSGIKYDQDLREG-----AGDLRAAFWLDLLESA-- 278

DB 4619 LRN-----SUHAATGLRLPATLIFDYPTPEALGYLR-----VELLREADD 4659

QY 278 GWGPRHDFKPPRTEDIDGVNWSA-----AGCMRNLYLKERAAAFRAADPEVEALR 330

DB 4660 GLDG-----REDDLRLVLAANVFAFKEAGVLDITLLGLADITGTEGCTDAETEARP 4710

QY 331 AARLDQALAEPTAADGL 346

DB 4711 AADDAELDALDISGL 4726

RESULT 14

ID Q9ZGA4 PRELIMINARY; PRT; 7576 AA.

AC Q9ZGA4;

DT 01-MAY-1999 (TRENBLrel. 10, Created)

DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)

DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)

DE FK506 POLYKETIDE SYNTHASE.

GN FKBB.

OS Streptomyces sp. MA6548.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MA6548;

RA MEDLINE; 98451508.

RA MOTAMEDI H., SHAFIEE A.;

RT "The biosynthetic gene cluster for the macrolactone ring of the

RT immunosuppressant FK506."

RL Eur. J. Biochem. 256:528-534(1998).

DR EMBL; AF082100; AAC68815.1; -.

DR HSP; P08659; 1LCI.

DR PROSITE; PS00455; AMP\_BINDING; 1.

DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; 4.

KW Transferase.

SQ SEQUENCE 7576 AA; 790105 MW; 42E87D19 CRC32;

Query Match 5.0%; Score 100.5; DB 2; Length 7576;

Best Local Similarity 24.7%; Pred. No. 4.5e-02;

Matches 92; Conservative 31; Mismatches 167; Indels 83; Gaps 19;

QY 25 PFGDTRPALDPVQRLAEAGVGTFFHDDLLIPFGASDTEREAHVKFRQALDAGTMT 84

DB 3203 PAAPAPLPADPVP-----EDVVVPSVARTPDGVRDV-----AGRLTALT 3242

QY 85 VPMA-----TTNLFTH-----PVFKAGAFANDR---AVRRYALRKTRINIDLAVEIG 129

DB 3243 APAAAIGHSLATRTAMRHRRAVVPARDAEAFARGEVEPGVVRGTADVTDTTRAVFVPPGOG 3302

QY 130 AKVYVWVGREGAESAAGKDVRAALDRMKEAFDILLGEYVTSQYDIRFAIEPKPNP--- 187

DB 3303 SQ-----WDGM-CAELLATEPVFAR--RLGECALAPYT--GWDLDDVIARRPGAPELD 3352

QY 187 RGDILLPTTIGHALAFIERLERPELYGVNPE--VGHEQ-----MAG-LNPHGIAQALW 236

DB 3353 RYVVQVPASFAMVALAELWR--AHGVAPAAVVGHSQGEVAAAACVAGVLTLDAAKVVWAL 3410

QY 237 AGKLFHIDLNGSGIKYDQDLREGAGDLRAAFWLDLLESAAGWEGPRHDFRPTREDID 296

DB 3411 RSRLVATERAGHGM-----VSVPPADFDAAAW-AGRLEVAAVNGFPAST----- 3454

QY 297 GWASAAGCMRNLYLKERAAAFRAADPEVQ-----BALRAARLDQALAEPT-AADGLQALLA 351

DB 3454 -VVAGNADAVEELLAAATPHARRIANDYASHTAHVESIRGALLDALADLTTPGAPEIPFFST 3512

QY 352 DRTAYEDEFVDAA 364

DB 3513 VDEAWLDRPADAA 3525

RESULT 15

ID Q9ZC00 PRELIMINARY; PRT; 373 AA.

AC Q9ZC00;

DT 01-MAY-1999 (TRENBLrel. 10, Created)

DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)

DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)

DE HYPOTHETICAL 40.7 KD PROTEIN.

GN SC186.22C.

OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RA SEEGER K.J., HARRIS D.;

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RA MEDLINE; 97000351.

RA REDENBACH M., KIESER H.M., DENAPATTE D., EICHNER A., CULLUM J.,

RA KINASHI H., HOPWOOD D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for

RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."

RL Mol. Microbiol. 21:77-96(1996).

DR EMBL; AL033505; CAA22050.1; -.

KW Hypothetical protein.

SQ SEQUENCE 373 AA; 40702 MW; AEA38C6C CRC32;

Query Match 4.8%; Score 97; DB 2; Length 373;

Best Local Similarity 22.4%; Pred. No. 10;

Matches 72; Conservative 29; Mismatches 108; Indels 112; Gaps 17;

QY 93 FTHPVFK-----AGAFANDRAVRYALRKTTIRNIDLAVELGAKVYVWVGREGSGAA 147

DB 5 FLUHYEHOGPWASVYVDSLRTEDTPHRELTAAAVAREL-----AESGADATC 55

QY 148 KDVRAALDRMKEAFDILLGEYVTSQYDIRFAIEPKPNPNEPRGDILLPTIGHALAFIERLER 207

DB 56 RAVREANDLERATD-----PHGRALFACAGQ-VVLDPALAR 91

```

Qy 208 PELXGVNPEVGEQAGLNFPHGIAQALWAGKLFHI-----DLNGQSGI-----KYD 254
Db 92 P-----PYGGTTADWA-PLPHVTPLLDLAGEDPVCVVAYIDRKGAD 131
Qy 255 QDLRFG-----AGDLRAAFNLVDLLESAGWEGPRHFDFKPPRTEDIDGVNA-----S 301
Db 132 FELRSALGSSDAGGVTRQWPVHRTSSADW-SERHFQLR-----VENTWEHNAETADA 184
Qy 302 AAGCMR-----NYLIL-----KERAAA-----FRADPEVCEALRAA---RLDQLAEPTRAADG 345
Db 185 LAVCOETGADLLILVGDRERRSVHERLPLRLQERVAEASRGAGSRLLDDEVEGLRDDH 244
Qy 346 LQALLADRTAYEDEFVDAAAR 366
Db 245 VRA-----RAREDLDRFLAAR 260

```

Search completed: March 22, 2000, 23:47:30  
Job time: 211 sec



Result No.	Query Match	Score	Length	DB	ID	Description
1	94.4	1916	386	1	XYLA_STROL	P15587 streptomyc
2	93.6	1900	388	1	XYLA_STRMR	P37031 streptomyc
3	93.0	1887.5	388	1	XYLA_STRVO	P09033 streptomyc
4	92.3	1872	387	1	XYLA_STRRU	P24300 streptomyc
5	91.8	1863	390	1	XYLA_STRAL	P24299 streptomyc
6	90.4	1834	388	1	XYLA_STRDI	P50910 streptomyc
7	75.6	1534.5	393	1	XYLA_STRDO	P23857 streptomyc
8	66.4	1346.5	393	1	XYLA_ACTMI	P12851 actinoplan
9	66.3	1344.5	394	1	XYLA_AMSP	P10654 ampullariel
10	62.9	1277	394	1	XYLA_ARTS7	P12070 arthrobacte
11	55.5	1127	387	1	XYLA_THEHT	P26997 thermus aqu
12	54.2	1100	387	1	XYLA_THECA	P56681 thermus aqu
13	16.4	333.5	449	1	XYLA_LACPE	P21938 lactobacill
14	16.2	329	448	1	XYLA_BACLI	P27832 bacillus li
15	16.0	324.5	435	1	XYLA_TETHA	O82845 tetragenoco
16	15.8	320	444	1	XYLA_THENE	P45687 thermotoga
17	15.7	318	439	1	XYLA_CLOTS	P29441 clostridium
18	15.5	314	449	1	XYLA_LACBR	P29443 lactobacill
19	15.4	312.5	439	1	XYLA_STAXY	P27157 staphylococ
20	15.3	310.5	439	1	XYLA_THETU	P3148 thermoanaer
21	15.3	310	439	1	XYLA_THESA	P30435 thermoanaer
22	14.9	301.5	438	1	XYLA_THEET	P22842 thermoanaer
23	14.8	299.5	445	1	XYLA_BACME	O08325 bacillus me
24	14.7	299	441	1	XYLA_BACSP	P54272 bacillus sp
25	14.7	299	441	1	XYLA_BACST	P54273 bacillus st
26	14.7	297.5	445	1	XYLA_BACSU	P04788 bacillus su
27	13.7	277.5	440	1	XYLA_ECOLI	P00944 escherichia
28	13.5	273.5	440	1	XYLA_KLEPN	P29442 klebsiella
29	13.1	269	479	1	XYLA_HORVU	Q40082 hordeum vul
30	11.8	239.5	439	1	XYLA_HAEIN	P44398 haemophilus
31	5.2	106	3491	1	TRQ1_SACER	Q03131 saccharopol
32	4.9	98.5	1061	1	TRQ4_ECOLI	P27189 escherichia
33	4.8	97	1381	1	VCAP_EBV	P03226 Epstein-bar
34	4.7	96	800	1	GUN_BAC31	P06564 bacillus sp

```
FT TURN 22 25
FT HELIX 36 40
FT TURN 41 42
FT STRAND 52 52
FT TURN 55 57
FT TURN 64 71
FT TURN 75 76
FT STRAND 84 84
FT TURN 101 102
FT TURN 104 105
FT HELIX 106 120
FT STRAND 125 127
FT TURN 130 131
FT TURN 141 142
FT TURN 144 149
FT HELIX 150 153
FT TURN 154 155
FT TURN 159 160
FT STRAND 177 179
FT TURN 197 198
FT HELIX 199 205
FT TURN 206 206
FT TURN 221 222
FT HELIX 233 239
FT STRAND 247 247
FT TURN 260 262
FT TURN 267 275
FT STRAND 283 284
FT TURN 295 296
FT TURN 298 300
FT HELIX 302 305
FT TURN 306 307
FT HELIX 308 311
FT TURN 312 320
FT HELIX 322 326
FT TURN 327 328
FT HELIX 343 345
FT TURN 346 347
FT TURN 362 363
FT HELIX 364 367
FT TURN 368 369
SQ SEQUENCE 386 AA; 42792 MW; 0F0B8308 CRC32;

Query Match 94.4%; Score 1916; DB 1; Length 386;
Best Local Similarity 94.3%; Pred. No. 6.1e-143;
Matches 364; Conservative 13; Mismatches 7; Indels 2; Gaps 2;

QY 2 NYQPTPEDRFTGLTVGWQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIPF 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 SQQTPEDRFTGLTVGWQGRDPFGDTRPALDPVEIVQRLAELGAGVTFHDDDLIPF 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 GASDTEREAVKRFQALDAGTGMVPMATNLTTHPVFKAGAFANDRAVRVYALRKTI 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 GSSDTERESHKRFQALDAGTGMVPMATNLTTHPVFKDGGFTANDRDRVYALRKTI 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 121 NIDLAVELGAKYVAVWGREGAESGAADVRAALDRMKEAFDLGELYTSQGYDIRFATE 180
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 121 NIDLAVELGAKYVAVWGREGAESGAADVRAALDRMKEAFDLGELYTSQGYDIRFAIE 180
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 181 PKPNEPRGDILLPTGHALAFIERLERPELYGVNPEVGEHQAGLNFPHGIAQALWAGKL 240
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 181 PKPNEPRGDILLPTGVHALAFIERLERPELYGVNPEVGEHQAGLNFPHGIAQALWAGKL 240
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 241 FHIIDNGSQGIYDQDLRFAGDGLRAAFWLVDLLSAGWEGPRHFDKPPRTEDIDGVWA 300
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 241 FHIIDNGSQGIYDQDLRFAGDGLRAAFWLVDLLSAGWEGPRHFDKPPRTEDIDGVWA 300
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 301 SAAGCMRNVLILKRAAFAADPEVQEAALRAARLDQAEPTAAGLQALLADRTAYEDFD 360
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 301 SAAGCMRNVLILKRAAFAADPEVQEAALRAARLDQAEPTAAGLQALLADRTAYEDFD 360
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 361 VD-AAARGMAFERLDQLAMDHLLGAR 385
```

```
Db 361 VDAARGMAFERLDQLAMDHLLGAR 386
|||||
RESULT 2
XYLA_STRMR
ID XYLA_STRMR STANDARD; PRT; 388 AA.
AC P37031;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE XYLOSE ISOMERASE (EC 5.3.1.5).
GN XYLA.
OS Streptomyces murinus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RA RASMUSSEN H., LA COUR T., NYBORG J., SCHUELEIN M.;
RT "Structure determination of glucose isomerase from Streptomyces
   murinus at 2.6-A resolution."
RL Acta Crystallogr. D 50:124-131(1994).
CC -|- FUNCTION: INVOLVED IN D-XYLOSE CATABOLISM.
CC -|- CATALYTIC ACTIVITY: D-XYLOSE = D-XYLULOSE.
CC -|- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
CC -|- SUBUNIT: HOMOTETRAMER.
CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -|- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
EMBL; AL0243; CAA00885.1; -.
PDB; 1DXI; 31-JAN-94.
DR PROSITE; PS00172; XYLOSE_ISOMERASE_1; 1.
DR PROSITE; PS00173; XYLOSE_ISOMERASE_2; 1.
PFAM; PF00259; Xylose_isom; 1.
KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium; 3D-structure.
FT ACT_SITE 54 54 BY SIMILARITY.
FT METAL 181 181 MAGNESIUM.
FT METAL 217 217 MAGNESIUM.
FT METAL 245 245 MAGNESIUM.
FT METAL 287 287 MAGNESIUM.
FT TURN 7 8
FT STRAND 11 14
FT HELIX 15 18
FT TURN 19 19
FT HELIX 36 46
FT TURN 47 47
FT STRAND 50 54
FT HELIX 55 58
FT HELIX 65 82
FT TURN 83 83
FT STRAND 85 85
FT STRAND 88 90
FT TURN 97 99
FT TURN 104 105
FT HELIX 109 127
FT TURN 128 130
FT STRAND 133 136
FT STRAND 142 143
FT HELIX 151 171
FT TURN 172 173
FT STRAND 178 180
FT STRAND 190 191
FT HELIX 196 203
FT TURN 204 205
FT TURN 209 211
FT HELIX 211
```

```

FT TURN 212 212
FT STRAND 214 214
FT STRAND 217 217
FT STRAND 218 222
FT TURN 223 225
FT HELIX 228 235
FT TURN 236 239
FT STRAND 244 246
FT STRAND 248 248
FT STRAND 258 258
FT TURN 260 261
FT HELIX 265 277
FT TURN 278 279
FT STRAND 284 286
FT TURN 292 293
FT HELIX 296 320
FT TURN 321 322
FT HELIX 324 332
FT TURN 333 334
FT HELIX 335 338
FT TURN 339 339
FT TURN 349 350
FT TURN 354 359
FT HELIX 362 366
FT TURN 367 367
FT HELIX 372 378
FT TURN 379 380
SQ SEQUENCE 388 AA; 42772 MW; F62EE677 CRC32;

```

```

Query Match 93.6%; Score 1900; DB 1; Length 388;
Best Local Similarity 92.8%; Pred. No. 1.le-141;
Matches 360; Conservative 14; Mismatches 12; Indels 2; Gaps 2;

QY 1 MNTQPTPEDFTFGLTWVGWQGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 59
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MSQPTPEDFTFGLTWVGWQGRDPFGDTRPALDPVETVORLAELGAYGVTFHDDDLIP 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 60 FGASDTEREAHVKKFRQALDGTMTVPMTNLTFTHPVFKAGFTANDRAVRRYALRKT 119
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 FGSDTERESHKFRQALDGTMTVPMTNLTFTHPVFKDGGFTANDROVRRYALRKT 120
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 120 RNTDLAVELGAKVYVANGREGAEGSAKDVRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 GNTDLAAELGAKTYVANGREGAEGSAKDVRAALDRMKEAFDILLGEYVTAQGYDLRFAI 180
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 180 EKPNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGK 239
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 EKPNEPRGDILLPTVGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGK 240
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 240 LFHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVW 299
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 LFHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVW 300
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 300 ASAGCMRNLYLILKERAARADPEVQEAALRAALDQAEPTAADGLQALLADRTAYEDF 359
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 ASAGCMRNLYLILKERAARADPEVQEAALRAALDQAEPTAADGLQALLADRTAYEDF 360
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 360 DVD-AAARGMAFERLDQALMDHLLGARG 386
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 DVDAARGMAFERLDQALMDHLLGARG 388
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

RESULT 3
XYLA_STRVO STANDARD; PRT; 388 AA.
AC P09033;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE XYLOSE ISOMERASE (EC 5.3.1.5).
GN XYLA.
OS Streptomyces violaceoniger.

```

```

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RX STRAIN-CBS 409-73;
MEDLINE; 89016633.
RA DROCOURT D., BEJAR S., CALMELS T., REYNES J.-P., TIRABY G.;
RT "Nucleotide sequence of the xylose isomerase gene from Streptomyces
violaceoniger.";
RL Nucleic Acids Res. 16:9337-9337(1988).
CC 1- FUNCTION: INVOLVED IN D-XYLOSE CATABOLISM.
CC 1- CATALYTIC ACTIVITY: D-XYLOSE = D-XYULOSE.
CC 1- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
CC 1- SUBUNIT: HOMOTETRAMER.
CC 1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC 1- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licenses@isb-sib.ch).
CC -----
DR EMBL; X12816; CAA31304.1; -
DR EMBL; M36269; AAA26839.1; -
DR PIR; S01436; ISSMXV.
DR HSP; P37031; 1DXI.
DR PROSITE; PS00172; XYLOSE ISOMERASE_1; 1.
DR PROSITE; PS00173; XYLOSE ISOMERASE_2; 1.
DR FRAM; PF00359; Xylose_isom; 1.
KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium.
FT INIT MET 0 0 BY SIMILARITY.
FT ACT_SITE 53 53 MAGNESIUM (BY SIMILARITY).
FT METAL 180 180 MAGNESIUM (BY SIMILARITY).
FT METAL 216 216 MAGNESIUM (BY SIMILARITY).
FT METAL 244 244 MAGNESIUM (BY SIMILARITY).
FT METAL 286 286 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 388 AA; 43013 MW; 37D69108 CRC32;

Query Match 93.0%; Score 1887.5; DB 1; Length 388;
Best Local Similarity 92.3%; Pred. No. 1e-140;
Matches 358; Conservative 15; Mismatches 12; Indels 3; Gaps 2;

QY 2 NYOPTPEDFTFGLTWVGWQGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 SFQPTPEDFTFGLTWVGWQGRDPFGDTRPALDPVETVORLAELGAYGVTFHDDDLIP 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 GASDTEREAHVKKFRQALDGTMTVPMTNLTFTHPVFKAGFTANDRAVRRYALRKTIR 120
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 GSSDTERESHKFRQALDGTMTVPMTNLTFTHPVFKDGGFTANDROVRRYALRKTIR 120
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 NIDLAVELGAKVYVANGREGAEGSAKDVRAALDRMKEAFDILLGEYVTSQGYDIRFAIE 180
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 NIDLAAELGAKTYVANGREGAEGSAKDVRAALDRMKEAFDILLGEYVTAQGYDLRFAIE 180
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 PKPNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGKL 240
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 PKPNEPRGDILLPTVGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGKL 240
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 241 FHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVWA 300
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 FHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGWEGPRHFDKPPRTEDIDGVWA 300
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 301 SAAGCMRNLYLILKERAARADPEVQEAALRAALDQAEPTAADGLQALLADRTAYEDFD 360
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 SAEGCMRNLYLILKERAARADPEVQEAALRAALDQAEPTAADGLQALLADRTAYEDFD 360
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 361 VDAARGMA--FERLDQALMDHLLGARG 386
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 VEAAAAARAAMPFERLDQALMDHLLGARG 388
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```



```
Db 61 GSSDSEEHVKKFRQALDDTGKVPMTNLFTHPVFKDGGFTANDRDVRRYALRKRTIR 120
QY 121 NIDLAVELGAKVYVANGREGAEGSAKADYRAALDRMKEAFDILLGEYVTSQGYDIRFAIE 180
Db 121 NIDLAVELGAETVANGREGAEGSAKADYRDALDRMKEAFDILLGEYVTSQGYDIRFAIE 180
QY 181 PKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNPPHGIQAALWAGKL 240
Db 181 PKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNPPHGIQAALWAGKL 240
QY 241 FHIDLNGOSGIKVDQDLRFAGAGDLRAAFWLVLDLESAGWEGPRHDFKPTPTEDIDGVA 300
Db 241 FHIDLNGQNGIKVDQDLRFAGAGDLRAAFWLVLDLESAGISGPRHDFKPTPTEDIDGVA 300
QY 301 SAAGCMRNILKERRAAFRADPEVQEAALRAALDQAEPTAADGLQALLADRTAYEDFD 360
Db 301 SAAGCMRNILKERRAAFRADPEVQEAALRASRLDELARPTAADGLQALLADRSAPFEFD 360
QY 361 VD-AAARGMAFERLDQALDHLHGARG 386
Db 361 VDAARGMAFERLDQALDHLHGARG 387

RESULT 5
XYLA_STRAL STANDARD; PRT; 390 AA.
AC P24299;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE XYLOSE ISOMERASE (EC 5.3.1.5).
GN XYLA.
OS Streptomyces albus 9.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
RX MEDLINE; 91197439.
RA DAUTER Z., TERRY H., WITZEL H., WILSON K.S.;
RT "Refinement of glucose isomerase from Streptomyces albus at 1.65 A
RT with data from an imaging plate."
RL Acta Crystallogr. B 46:833-845(1990).
[2]
RN RP
RX X-RAY CRYSTALLOGRAPHY.
RX MEDLINE; 89211422.
RA DAUTER Z., DAUTER M., HEMKER J., WITZEL H., WILSON K.S.;
RT "Crystallisation and preliminary analysis of glucose isomerase from
RT Streptomyces albus."
RL FEBS Lett. 247:1-8(1989).
CC -!- FUNCTION: INVOLVED IN D-XYLOSE CATABOLISM.
CC -!- CATALYTIC ACTIVITY: D-XYLOSE -> D-XYLULOSE.
CC -!- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.
DR PDB; 6XLA; 15-JUL-92.
DR PROSITE; PS00172; XYLOSE ISOMERASE_1; 1.
DR PROSITE; PS00173; XYLOSE ISOMERASE_2; 1.
DR PFAM; PF00359; Xylose_isom; 1.
KW isomerase; Pentose shunt; Xylose metabolism; Magnesium; 3D-structure.
FT ACT_SITE 53 53 BY SIMILARITY.
FT METAL 180 180 MAGNESIUM (BY SIMILARITY).
FT METAL 216 216 MAGNESIUM (BY SIMILARITY).
FT METAL 244 244 MAGNESIUM (BY SIMILARITY).
FT METAL 286 286 MAGNESIUM (BY SIMILARITY).
FT HELIX 6 8
FT STRAND 10 13
FT HELIX 14 17
FT TURN 18 18
FT STRAND 23 23
FT TURN 24 25
FT STRAND 26 26
```

```
FT HELIX 35 45
FT TURN 46 46
FT STRAND 49 53
FT HELIX 54 57
FT TURN 60 61
FT HELIX 64 81
FT TURN 82 82
FT STRAND 84 84
FT STRAND 87 89
FT HELIX 96 98
FT TURN 99 100
FT TURN 103 104
FT HELIX 108 127
FT TURN 128 129
FT STRAND 132 135
FT TURN 138 139
FT STRAND 141 142
FT HELIX 145 147
FT HELIX 150 171
FT TURN 172 172
FT STRAND 176 179
FT STRAND 189 190
FT HELIX 195 202
FT TURN 203 204
FT HELIX 208 210
FT STRAND 211 213
FT STRAND 216 216
FT HELIX 217 221
FT TURN 222 224
FT HELIX 227 236
FT TURN 237 238
FT STRAND 240 245
FT STRAND 244 245
FT STRAND 247 247
FT STRAND 257 257
FT TURN 259 260
FT HELIX 264 276
FT TURN 277 278
FT STRAND 283 285
FT TURN 291 292
FT HELIX 295 321
FT HELIX 323 331
FT TURN 332 333
FT HELIX 334 337
FT TURN 338 338
FT TURN 343 344
FT HELIX 346 351
FT HELIX 353 355
FT TURN 356 358
FT HELIX 361 366
FT HELIX 371 383
FT TURN 384 384
SQ SEQUENCE 390 AA; 43289 MW; 88ABE836 CRC32;

Query Match 91.8%; Score 1863; DB 1; Length 390;
Best Local Similarity 91.7%; Pred. No. 8.7e-139;
Matches 355; Conservative 15; Mismatches 15; Indels 2; Gaps 2;

QY 2 NYOPTPEDRFTGLWTVGNQGRDPFGDATERPALDPVD-VQRLAELGAYGVTFHDDDLIPF 60
Db 1 NYOPTPEDRFTGLWTVGNQGRDPFGDATERPALDPVESVRRLAELGAHGVTFFHDDDLIPF 60
QY 61 GASDTEREAHVKKFRQALDGTGTVPMATNLFTHPVFKAGAFANDRAVRRYALRKRTIR 120
Db 61 GSSDSERYEHVKKFRQALDDTGKVPMTNLFTHPVFKDGGFTANDRDVRRYALRKRTIR 120
QY 121 NIDLAVELGAKVYVANGREGAEGSAKADYRAALDRMKEAFDILLGEYVTSQGYDIRFAIE 180
Db 121 NIDLAVELGAETVANGREGAEGSAKADYRDALDRMKEAFDILLGEYVTSQGYDIRFAIE 180
QY 181 PKNEPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNPPHGIQAALWAGKL 240
```

```

Db 181 PKPNEPRGDILLPTVYCHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 240
Qy 241 FHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLLSAGWEGPRHDFKPPRTEDIDGVWA 300
Db 241 FHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLLSAGWEGPRHDFKPPRTEDIDGVWA 300
Qy 301 SAAGCMRNLYLILKRAAAFRADPEVQEAALRAARLDQAEPTAADGLQALLADRTAYEDFD 360
Db 301 SAAGCMRNLYLILKRAAAFRADPEVQEAALRAARLDQAEPTAADGLQALLADRTAYEDFD 360
Qy 361 VD-AAARGMAFERLDQALMDHLLGARG 386
Db 361 VDAAAARGMAFERLDQALMDHLLGARG 387

RESULT 6
XYLA_STRDI STANDARD; PRT; 388 AA.
AC P50910;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE XYLOSE ISOMERASE (EC 5.3.1.5).
GN XYLA.
OS Streptomyces diastaticus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NO. 7 / M1033;
RX MEDLINE; 95101954.
RA WANG Y., HUANG Z., DAI X., LIU J., CUI T., NIU L., WANG C., XU X.;
RT "The sequence of xylose isomerase gene from Streptomyces diastaticus No. 7 M1033.";
RL Chin. J. Biotechnol. 10:97-103(1994).
CC -|- FUNCTION: INVOLVED IN D-XYLOSE CATABOLISM.
CC -|- CATALYTIC ACTIVITY: D-XYLOSE -> D-XYLULOSE.
CC -|- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
CC -|- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -|- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S73809; AAB32873.1; -.
DR HSP; P24300; 2XIS.
DR PROSITE; PS00172; XYLOSE_ISOMERASE_1; 1.
DR PROSITE; PS00173; XYLOSE_ISOMERASE_2; 1.
DR PFAM; PF00259; Xylose_isom; 1.
KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium.
FT ACT_SITE 54 54 BY SIMILARITY.
FT METAL 181 181 MAGNESIUM (BY SIMILARITY).
FT METAL 217 217 MAGNESIUM (BY SIMILARITY).
FT METAL 245 245 MAGNESIUM (BY SIMILARITY).
FT METAL 287 287 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 388 AA; 42675 MW; 5E05C770 CRC32;

```

Query Match 90.4%; Score 1834; DB 1; Length 388;  
 Best Local Similarity 89.9%; Pred. No. 1.6e-136;  
 Matches 349; Conservative 16; Mismatches 21; Indels 2; Gaps 2;

```

Qy 1 MNQPTPEDRFTFGLTWVGWQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 59
Db 1 MSIQTPEDKFTFGLTWVGWQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 60
Qy 60 FGASDTEREAHVKFRQALDATGMTVPMTNLTNTHPVKAGAFANDRAVRRYALRKTI 119

```

```

Db 61 FGATDSERAEHIKFRGCLDETGMKVPMTNLTNTHPVKAGAFANDRAVRRYALRKTI 120
Qy 120 RNIDLAVELGAKVYVANGREGAAGAAKADYRAALDRMKKEAFDILLGEYVTSQGYDIRFAI 179
Db 120 RNIDLAVELGAKVYVANGREGAAGAAKADYRAALDRMKKEAFDILLGEYVTSQGYDIRFAI 179
Qy 180 EPKNEPRGDILLPTVYCHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 239
Db 180 EPKNEPRGDILLPTVYCHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 239
Qy 240 LFIHDLNGSGIKYDQDLRFAGDGLRAAFWLVLLSAGWEGPRHDFKPPRTEDIDGVW 299
Db 240 LFIHDLNGSGIKYDQDLRFAGDGLRAAFWLVLLSAGWEGPRHDFKPPRTEDIDGVW 299
Qy 300 ASAGCMRNLYLILKRAAAFRADPEVQEAALRAARLDQAEPTAADGLQALLADRTAYEDFD 359
Db 300 ASAGCMRNLYLILKRAAAFRADPEVQEAALRAARLDQAEPTAADGLQALLADRTAYEDFD 359
Qy 360 DVD-AAARGMAFERLDQALMDHLLGARG 386
Db 360 DVD-AAARGMAFERLDQALMDHLLGARG 386
Qy 361 VDAAAARGMAFERLDQALMDHLLGARG 387
Db 361 VDAAAARGMAFERLDQALMDHLLGARG 387

RESULT 7
XYLA_STRRO STANDARD; PRT; 393 AA.
AC P22857;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE XYLOSE ISOMERASE (EC 5.3.1.5).
GN XYLA.
OS Streptomyces rochei (Streptomyces parvullus).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S-41;
RX MEDLINE; 91136786.
RA KIKUCHI T., ITOH Y., KASUMI T., FUKAZAWA C.;
RT "Molecular cloning of the xylA gene encoding xylose isomerase from Streptomyces griseofuscus S-41: primary structure of the gene and its product.";
RL Agric. Biol. Chem. 54:2469-2472(1990).
CC -|- FUNCTION: INVOLVED IN D-XYLOSE CATABOLISM.
CC -|- CATALYTIC ACTIVITY: D-XYLOSE -> D-XYLULOSE.
CC -|- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
CC -|- SUBUNIT: HOMOTETRAMER.
CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -|- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.
DR HSP; P37031; 1DXI.
DR PROSITE; PS00172; XYLOSE_ISOMERASE_1; 1.
DR PROSITE; PS00173; XYLOSE_ISOMERASE_2; 1.
DR PFAM; PF00259; Xylose_isom; 1.
KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium.
FT INIT_MET 0 0 BY SIMILARITY.
FT ACT_SITE 53 53 MAGNESIUM (BY SIMILARITY).
FT METAL 179 179 MAGNESIUM (BY SIMILARITY).
FT METAL 215 215 MAGNESIUM (BY SIMILARITY).
FT METAL 243 243 MAGNESIUM (BY SIMILARITY).
FT METAL 284 284 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 393 AA; 43513 MW; 941FD6DE CRC32;

```

Query Match 75.6%; Score 1534.5; DB 1; Length 393;  
 Best Local Similarity 79.2%; Pred. No. 4.6e-113;  
 Matches 305; Conservative 21; Mismatches 24; Indels 35; Gaps 8;

```

Qy 2 NYQPTPEDRFTFGLTWVGWQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 60
Db 1 SFQPTPEDKFTFGLTWVGWQGRDPFGDTRPALDPVD-VQRLAELGAYGVTFHDDDLIP 60

```

QY 61 GASDTEREAKVHKFRQALDQATGVTVPMTATNLTTHPVFKAGATANDRAVRVYALRKRTIR 120  
 DB 61 GSSDTERESHKFRQALDQATGVTVPMTATNLTTHPVFK-DREFTANDRDVRAVAVRKRTIR 119  
 QY 121 NIDLAELGAKTVVANGREGSAGKADVRALDRKKEAFDILGGEVTSOGYDIFRAIE 180  
 DB 120 NIDLAELGAKTVVANGREGSAGKADVRALDRKKEAFDILGGEVTSOGYDIFRAIE 179  
 QY 181 PKPNEPRGDILLPTIGHALAFIERLERPELYGNPEVGHQMAGLNFPHGIAQALWAGKL 240  
 DB 180 PKPNEPRGDILLPTIGHALAFIERLERPELYGNPEVGHQMAGLNFPHGIAQALWAGKL 239  
 QY 241 FHDNLGQSGIKYDQDL---RFGAGDLRAFWLVDLLESAGWEGPRHDFKPPRTEDIDG 297  
 DB 240 FHDNLGQSGIKYDQCGRRRPPAGV---LVVDLLESAGYEGPRHDFKPPRTEDFDG 295  
 QY 298 VWASAGCHRNLYLILKE-----RAAFRADPEVQALRAARLDQLAEPTAAD 344  
 DB 296 VWASAGCHRNLYLILKE-----RAAFRADPEVQALRAARLDQLAEPTAAD 343  
 QY 345 GLOALLADRTAYEDFDVD-AAARGM 368  
 DB 344 GLEALLADRTAFEDFDEVAARGM 368

RESULT 8  
 ID XYLACTMI STANDARD; PRT; 393 AA.  
 AC P12851;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE XYLOSE ISOMERASE (EC 5.3.1.5).  
 GN XYLA OR XI.  
 OS Actinoplanes missouriensis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Micromonosporineae; Micromonosporaceae;  
 OC Actinoplanes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DSM 43046;  
 RX MEDLINE; 90016811.  
 RA AMORE R., HOLLENBERG C.P.;  
 RT "xylose isomerase from Actinoplanes missouriensis: primary structure  
 of the gene and the protein."  
 RL Nucleic Acids Res. 17:7515-7515(1989).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE; 89184498.  
 RA REY F., JENKINS J., JANIN J., LASTERS I., ALARD P., CLAESSENS M.,  
 RA MATTHYSSENS G., WODAK S.G.;  
 RT "Structural analysis of the 2.8 A model of xylose isomerase from  
 Actinoplanes missouriensis."  
 RL Proteins 4:165-172(1988).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
 RX MEDLINE; 92304915.  
 RA JENKINS J., JANIN J., REY F., CHIADMI M., VAN TILBEURGH H.,  
 RA LASTERS I., DE MAEYER M., VAN BELLE D., WODAK S.J., LAUWEREYS M.,  
 RA STANSSENS P., MRABET N.T., SNAUWAERT J., MATTHYSSENS G.,  
 RA LAMBEIR A.-M.;  
 RT "Protein engineering of xylose (glucose) isomerase from Actinoplanes  
 missouriensis. 1. Crystallography and site-directed mutagenesis of  
 metal binding sites."  
 RL Biochemistry 31:5449-5458(1992).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (4.1 ANGSTROMS).  
 RA RAMIN M., SHEPARD W., FOURME R., KAHN R.;  
 RL Submitted (JUN-1998) to the PDB data bank.  
 RN [5]  
 RP MUTAGENESIS.  
 RX MEDLINE; 92304916.

RA LAMBEIR A.-M., LAUWEREYS M., STANSSENS P., MRABET N.T., SNAUWAERT J.,  
 VAN TILBEURGH H., MATTHYSSENS G., LASTERS I., DE MAEYER M.,  
 WODAK S.J., JENKINS J., CHIADMI M., JANIN J.;  
 RT "Protein engineering of xylose (glucose) isomerase from Actinoplanes  
 missouriensis. 2. Site-directed mutagenesis of the xylose binding  
 site."  
 RL Biochemistry 31:5459-5466(1992).  
 RN [6]  
 RP MUTAGENESIS.  
 RX MEDLINE; 92304917.  
 RA VAN TILBEURGH H., JENKINS J., CHIADMI M., JANIN J., WODAK S.J.,  
 MRABET N.T., LAMBEIR A.-M.;  
 RT "Protein engineering of xylose (glucose) isomerase from Actinoplanes  
 missouriensis. 3. Changing metal specificity and the pH profile by  
 site-directed mutagenesis."  
 RL Biochemistry 31:5467-5472(1992).  
 CC -!- CATALYTIC ACTIVITY: D-XYLOSE = D-XYLULOSE.  
 CC -!- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.  
 CC -!- SUBUNIT: HOMOTETRAMER.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X16042; CAA34164.1;  
 DR EMBL; AL0241; CAA00884.1;  
 DR PIR; S05998; ISMAXM.  
 DR PDB; 1XIN; 15-JUL-93.  
 DR PDB; 2XIN; 15-JUL-93.  
 DR PDB; 3XIN; 15-JUL-93.  
 DR PDB; 5XIN; 15-JUL-93.  
 DR PDB; 1XIM; 15-APR-93.  
 DR PDB; 2XIM; 15-APR-93.  
 DR PDB; 3XIM; 15-APR-93.  
 DR PDB; 4XIM; 15-JUL-93.  
 DR PDB; 5XIM; 15-JUL-93.  
 DR PDB; 6XIM; 15-JUL-93.  
 DR PDB; 7XIM; 15-JUL-93.  
 DR PDB; 8XIM; 15-JUL-93.  
 DR PDB; 9XIM; 15-JUL-93.  
 DR PDB; 1BHW; 18-NOV-98.  
 DR PROSITE; PS00172; XYLOSE ISOMERASE\_1; 1.  
 DR PROSITE; PS00173; XYLOSE ISOMERASE\_2; 1.  
 DR PFW; PF00259; Xylose\_isom; 1.  
 KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium; 3D-structure.  
 FT INIT MET 0 0  
 FT ACT SITE 53 53  
 FT METAL 180 180  
 FT METAL 216 216  
 FT METAL 244 244  
 FT METAL 291 291  
 FT METAL 6 8  
 FT HELIX 10 13  
 FT STRAND 14 17  
 FT TURN 18 18  
 FT STRAND 23 23  
 FT TURN 24 25  
 FT STRAND 26 26  
 FT TURN 35 45  
 FT HELIX 46 46  
 FT TURN 49 51  
 FT STRAND 53 53  
 FT TURN 54 57  
 FT TURN 60 61  
 FT HELIX 64 81  
 FT TURN 82 82  
 FT STRAND 84 84





2.

```

FT TURN 279 280
FT STRAND 289 291
FT TURN 297 298
FT TURN 301 327
FT HELIX 301 327
FT TURN 329 337
FT TURN 338 341
FT HELIX 342 344
FT TURN 350 351
FT HELIX 354 359
FT TURN 361 366
FT HELIX 369 372
FT TURN 373 374
FT HELIX 379 390
FT TURN 391 392
SQ SEQUENCE 394 AA: 43160 MW; BCB86DC2 CRC32;

Query Match 62.9%; Score 1277; DB 1; Length 394;
Best Local Similarity 62.5%; Pred. No. 6.6e-93;
Matches 245; Conservative 52; Mismatches 85; Indels 10; Gaps 4;

QY 4 QPTPEDRTFGLTVMGWRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIPFGA 62
DB 3 QPTPADHFTFGLTVMGWRDPFGDTRPALDPVD-EAVHKLAEGLGAYGTFHDDDLIPFGA 62
QY 63 SDTEREAHVKFRQALDQATGTMVPMATNLTTPVFKAGAFANDRAVRRYALRKTIIRNI 122
DB 63 TEAREKILGDFNQALKTGLKVPMTNLTSPHVPKDGFTSNDRSIRRFALAKVLHNI 122
QY 123 DLAVELGAKVYVANGREGAESGAADVRAALDRMKEAFDGLLGEYVTSQGYDIRFAIEPK 182
DB 123 DLAAEMGAETVMMGREGSEYDGSKLAAALDRMREGVDTAAGYIKDKGVNLRALIEPK 182
QY 183 PNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGHQMGAGLNFPHGIAQALWAGKLFH 242
DB 183 PNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGHQMGAGLNFPHGIAQALWAGKLFH 242
QY 243 IDLNQSGIKYDQDLRFAGAGDLRAAFWLVLDLES- - - - -GWGPRHFDKPPRTEDID 296
DB 243 IDLNQSGIKYDQDLRFAGAGDLRAAFWLVLDLES- - - - -GWGPRHFDKPPRTEDID 296
QY 297 GWMAAAGCMRNYLILKERAARAFADPEVQEAALRAARLDQALAEPT- -AAGLQALLADRT 354
DB 303 GWMSAKAMSMYLLKERAARAFADPEVQEAALRAARLDQALAEPT- -AAGLQALLADRT 354
QY 355 AYEDFDVDAAR-RGNAFERLDQLANDHLLGAR 385
DB 363 SFAGFDAAEAARNAFIRLNQALAEHLIGSR 394

RESULT 11
XYLA_THETH STANDARD; PRT; 387 AA.
AC P26997;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE XYLOSE ISOMERASE (EC 5.3.1.5).
GN XYLA.
OS Thermus aquaticus (subsp. thermophilus).
OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
RP SEQUENCE FROM N.A.
RX MEDLINE; 91216978.
RA DEKKER K., YAMAGATA H., SAKAGUCHI K., UDAKA S.;
RT "Xylose (glucose) isomerase gene from the thermophile Thermus
RT thermophilus: cloning, sequencing, and comparison with other
RT thermostable xylose isomerases."
RL J. Bacteriol. 173:3078-3083(1991).
RN (2)
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RC STRAIN=HB8;
RX MEDLINE; 99264393.

CHANG C., PARK B.C., LEE D.-S., SUH S.W.;
"Crystal structures of thermostable xylose isomerases from Thermus
calidophilus and Thermus thermophilus: possible structural
determinants of thermostability."
J. Mol. Biol. 288:623-634(1999).
CC -1- CATALYTIC ACTIVITY: D-XYLOSE = D-XYLULOSE.
CC -1- COFACTOR: MAGNESIUM IS NECESSARY FOR THE ACTIVITY.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90256; BAA14301.1; -.
DR PIR; A39404; A39404.
DR PDB; 1BX8; 16-FEB-99.
DR PROSITE; PS00174; XYLOSE-ISOMERASE_1; 1.
DR PROSITE; PS00173; XYLOSE-ISOMERASE_2; 1.
DR PFAM; PF00259; Xylose_isom. 1.
KW Isomerase; Pentose shunt; Xylose metabolism; Magnesium; 3D-structure.
FT ACT_SITE 53 53 BY SIMILARITY.
FT METAL 180 180 MAGNESIUM (BY SIMILARITY).
FT METAL 216 216 MAGNESIUM (BY SIMILARITY).
FT METAL 244 244 MAGNESIUM (BY SIMILARITY).
FT METAL 286 286 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 387 AA; 43906 MW; B72CCEC7 CRC32;

Query Match 55.5%; Score 1127; DB 1; Length 387;
Best Local Similarity 58.5%; Pred. No. 3.6e-81;
Matches 226; Conservative 43; Mismatches 109; Indels 10; Gaps 3;

QY 3 YQPTPEDRTFGLTVMGWRDPFGDTRPALDPV-DVORLAELGAYGVTFHDDDLIPFG 61
DB 2 YEPKPEHRFTFGLTVMGWRDPFGDTRPALDPV-VYVHKLAEGLGAYGVNLDLIPRG 61
QY 62 ASDTEREAHVKFRQALDQATGTMVPMATNLTTPVFKAGAFANDRAVRRYALRKTIIRN 121
DB 62 TPQERQIVRFRKALDETGLKVPMTNLTSPHVPKDGFTSNDRSIRRFALAKVLHNI 121
QY 122 DLAVELGAKVYVANGREGAESGAADVRAALDRMKEAFDGLLGEYVTSQGYDIRFAIEP 181
DB 122 MDLGAELGAEIYVYVANGREGAEVATGKARKVMDVWRVREALNFMAYAEADQGYRFALEP 181
QY 182 KPNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGHQMGAGLNFPHGIAQALWAGKLF 241
DB 182 KPNEPRGDIYFATVGMALFHTLDRPERFGLNPEFHTMAGLNFPHVAVQAADAGKLF 241
QY 242 HIDLNGSGIKYDQDLRFAGAGDLRAAFWLVLDLESAGWEGPRHFDKPPRTEDIDGWAS 301
DB 242 HIDLNDQMSRFDDDLRFGSENKAAFLVDLLESSGQVGRHFDHALRTDEGEVWAF 301
QY 302 AAGCMRNYLILKERAARAFADPEVQEAALRAARLDQ- - - - -LAETPAADGLQALLADRTAY 356
DB 302 ARGCMRTYILKERAARAFADPEVQEAALRAARLDQ- - - - -LAETPAADGLQALLADRTAY 356
QY 357 EDFDVAARAGMAFERLDQLANDHLLGAR 386
DB 362 E- - - - -AKRRRGYALERLDQLAVEYLLGVRG 387

RESULT 12
XYLA_THETH STANDARD; PRT; 387 AA.
AC P56681;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
```

RESULT 13



[illegible]

Search completed: March 22, 2000, 23:46:36  
Job time: 187 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2000, 23:31:44 ; Search time 34.38 Seconds  
(without alignments)  
529.592 Million cell updates/sec

Title: US-09-383-318-2

Perfect score: 2029  
Sequence: 1 MNYQPTPEDRFTGLWTGV.....GMAFERLDQLAMHLLGARG 386

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 142080 seqs, 47169319 residues

Total number of hits satisfying chosen parameters: 142080

Minimum DB seq length: 0  
Maximum DB seq length: 1000000  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : PIR\_62:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1892.5	93.3	389	1 ISSMXV	xylose isomerase (
2	1877	92.5	388	1 B41339	xylose isomerase (
3	1872.5	92.3	389	1 S28986	xylose isomerase (
4	1831	90.2	388	2 JC1031	xylose isomerase (
5	1539.5	75.9	394	1 ISSMXR	xylose isomerase (
6	1353.5	66.7	394	1 ISWAXA	xylose isomerase (
7	1351.5	66.6	394	1 ISWAXM	xylose isomerase (
8	1282	63.2	395	1 S16214	xylose isomerase (
9	1127	55.5	387	2 A39404	xylose isomerase (
10	333.5	16.4	449	2 S18561	xylose isomerase (
11	318	15.7	439	2 I40806	xylose isomerase (
12	317	15.6	444	2 A72225	xylose isomerase (
13	314	15.5	449	2 JC1137	xylose isomerase (
14	312.5	15.4	439	2 S16530	xylose isomerase (
15	310.5	15.3	439	1 ISCLXM	xylose isomerase (
16	301.5	14.9	438	1 ISCLXH	xylose isomerase (
17	297.5	14.7	445	1 ISBSXS	xylose isomerase (
18	277.5	13.7	440	1 ISBCK1	xylose isomerase (
19	273.5	13.5	440	1 ISKXB	xylose isomerase (
20	239.5	11.8	439	1 ISHIX	xylose isomerase (
21	235	11.6	479	2 S63466	xylose isomerase (
22	194	9.6	34	2 PS0335	xylose isomerase (
23	179	8.8	34	2 PS0336	sugar isomerase -
24	138.5	6.8	383	2 A72299	conserved hypothet
25	107	5.3	233	2 H69130	trac-1 protein - E
26	98.5	4.9	1061	2 S37667	major capsid prote
27	97	4.8	1381	1 Q0BE45	hypothetical prote
28	96	4.7	415	1 S75766	cellulase (EC 3.2.
29	96	4.7	800	2 A29003	
30	95.5	4.7	367	2 S66502	potassium channel

31	95	4.7	270	2 F72381	D-tagatose 3-epime
32	94	4.6	287	2 S77309	hypothetical prote
33	94	4.6	337	2 G75160	acetyl ornithine d
34	94	4.6	3591	1 S21010	filamentous hemag
35	93.5	4.6	480	2 D70600	probable polynucle
36	93.5	4.6	1404	2 E38788	hypothetical prote
37	93	4.6	588	2 A25902	65K antigen - Myco
38	93	4.6	822	2 JT0611	cellulase (EC 3.2.
39	92.5	4.6	367	2 A53131	alpha-dendrotoxin-
40	92.5	4.6	1486	2 C64832	cell division prot
41	92	4.5	540	2 A26950	groEL2 protein - M
42	92	4.5	540	2 A43509	65K antigen mbaA -
43	91.5	4.5	392	2 S15295	nolC protein - Rhl
44	91.5	4.5	482	2 S08384	modulation protein
45	91.5	4.5	482	2 S10133	

ALIGNMENTS

RESULT 1  
ISSMXV  
xylose isomerase (EC 5.3.1.5) - Streptomyces violaceusniger  
N:Alternate names: xylose ketol-isomerase  
C:Species: Streptomyces violaceusniger  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 18-Jun-1999  
C:Accession: S01436  
R:Drocourt, D.; Bejar, S.; Calmels, T.; Reynes, J.P.; Tiraby, G.  
Nucleic Acids Res. 16, 9337, 1988  
A:Title: Nucleotide sequence of the xylose isomerase gene from Streptomyces violaceo  
A:Reference number: S01436; MUID:89016633  
A:Accession: S01436  
A:Molecule type: DNA  
A:Residues: 1-389 <DRO>  
A:Cross-references: EMBL:X12816; NID:948004; PIDN:CAA31304.1; PID:948005  
A:Note: part of this sequence, including the amino end of the mature protein, was con  
C:Genetics:  
A:Gene: xyla  
C:Complex: homotetramer  
C:Function:  
A:Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopento  
A:Pathway: xylose metabolism  
C:Superfamily: xylose isomerase  
C:Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; xylose  
F:2-389/Product: xylose isomerase #status predicted <MAT>  
F:54,57/Active site: His, Asp #status predicted  
F:90,183/Binding site: substrate (Thr, Lys) #status predicted  
F:181,217,245,287/Binding site: magnesium (Glu, Glu, Asp, Asp) #status predicted  
F:217,220,255,257/Binding site: magnesium (Glu, His, Asp, Asp) #status predicted

Query Match	93.3%	Score	1892.5	DB 1	Length	389			
Best Local Similarity	92.3%	Pred. No.	3e-137						
Matches	359	Conservative	15	Mismatches	12	Indels	3	Gaps	2
QY	1	MNYQPTPEDRFTGLWTGVQGRDPFGDATRPALDPVD-VQRLAELGAYGVTFHDDDLIP	59						
Db	1	MSFQPTPEKFTGLWTGVQGRDPFGDATRPALDPVETVQRLAELGAYGVTFHDDDLIP	60						
QY	60	FGASDTEREAHVFRQALDATGCTVPMATNTLFTHPVFKAGAFANDRAVRRYALRKTI	119						
Db	61	FGSDTERESHIRFQALDATGCTVPMATNTLFTHPVFKDGGTANDRDVRRYALRKTI	120						
QY	120	RNIDLAVELGAKYVAVGREGAESGAADVRAALDRMKEAFDLLGEYVTSQGYDIRFAI	179						
Db	121	RNIDLAELGAKTYAVGREGAESGKADVRAALDRMKEAFDLLGEYVTAQGYDLRFAI	180						
QY	180	EPKPEPRGDILLPTIGHALAFIERLERPELYGVNPEVGEQVAGLNFPFHGIAQALWAGK	239						
Db	181	EPKPEPRGDILLPTVGHALAFIERLERPELYGVNPEVGEQVAGLNFPFHGIAQALWAGK	240						
QY	240	LPHIDLNGQSGIKYDQDLRFAGAGDLRAAFWLVDLLSAGWEGPRHDFKPRPTEDIDGVW	299						

Db 241 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLESAGYEGPRHDFKPPRTEDFDGVM 300  
Qy 300 ASAGCMRNLYLKERAAFRADPEVQALRAALDQALAEPTAADGLQALLADRTAYEDF 359  
Db 301 ASAGCMRNLYLKERAAFRADPEVQALRAALDQALAEPTAADGLQALLADRTAYEDF 360  
Qy 360 DVDAARAGMA--FERLDQALMDHLLGARG 386  
Db 361 DVDAARAGMA--FERLDQALMDHLLGARG 389  
RESULT 2  
B41339  
xylose isomerase (EC 5.3.1.5) - Streptomyces rubiginosus  
C:Species: Streptomyces rubiginosus  
C:Date: 28-May-1992 #sequence\_revision 06-Dec-1996 #text\_change 18-Jun-1999  
C:Accession: B41339  
R:Wong, H.C.; Ting, Y.; Lin, H.C.; Reichert, F.; Myambo, K.; Watt, K.W.K.; Toy, P.L.; D  
J. Bacteriol. 173, 6849-6858, 1991  
A:Title: Genetic organization and regulation of the xylose degradation genes in Streptom  
A:Reference number: A41339; MUID:92041569  
A:Accession: B41339  
A:Molecule type: DNA  
A:Residues: 1-388 <WON>  
A:Cross-references: GB:M73789; NID:g153532; PIDN:AAA26838.1; PID:g153534  
A:Note: parts of this sequence, including the amino and carboxyl ends of the mature pro  
R:Carroll, H.L.; Glusker, J.P.  
submitted to the Brookhaven Protein Data Bank, October 1990  
A:Reference number: A50791; PDB:7XIA  
A:Contents: annotation: X-ray crystallography, 1.9 angstroms, residues 1-40, 'Q', 42-387  
R:Carroll, H.L.; Glusker, J.P.  
submitted to the Brookhaven Protein Data Bank, October 1990  
A:Reference number: A50808; PDB:8XIA  
A:Contents: annotation: X-ray crystallography, 1.9 angstroms, with D-xylose, residues 1-  
R:Carroll, H.L.; Glusker, J.P.; Burger, V.; Manfre, F.; Tritsch, D.; Biellmann, J.F.  
Proc. Natl. Acad. Sci. U.S.A. 86, 4440-4444, 1989  
A:Title: X-ray analysis of D-xylose isomerase at 1.9 angstroms: native enzyme in complex  
A:Reference number: A58422; MUID:89282788  
A:Contents: annotation: X-ray crystallography, 1.9 angstroms  
R:Whitlow, M.; Howard, A.J.  
submitted to the Brookhaven Protein Data Bank, March 1991  
A:Reference number: A50703; PDB:4XIS  
A:Contents: annotation: X-ray crystallography, 1.60 angstroms, with xylose and manganese  
R:Whitlow, M.; Howard, A.J.; Finzel, B.C.; Poulos, T.L.; Winborne, E.; Gilliland, G.L.  
Proteins 9, 153-173, 1991  
A:Title: A metal-mediated hydride shift mechanism for xylose isomerase based on the 1.6  
A:Reference number: A58421; MUID:91172741  
A:Contents: annotation: X-ray crystallography, 1.60 angstroms  
C:Complex: homotetramer  
C:Function:  
A:Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopentose  
A:Pathway: xylose metabolism  
C:Superfamily: xylose isomerase  
C:Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; xylose me  
F:1-388/Product: xylose isomerase #status experimental <MAT>  
F:54.57/Active site: His, Asp #status experimental  
F:90.183/Binding site: substrate (Thr, Lys) #status experimental  
F:181,217,245,287/Binding site: magnesium (Glu, Glu, Asp, Asp) #status experimental  
F:217,220,255,257/Binding site: magnesium (Glu, His, Asp, Asp)

Query Match 92.5%; Score 1877; DB 1; Length 388;  
Best Local Similarity 92.3%; Pred. No. 4.5e-136;  
Matches 358; Conservative 14; Mismatches 14; Indels 2; Gaps 2;  
Qy 1 MNQYPTPEDRTFFGLTWVGWQGRDPFGDTRPALDPVD-VORLAEIGAYGVTFHDDDLIP 59  
Db 1 MNQYPTPEDRTFFGLTWVGWQGRDPFGDTRRALDPVESVRLAEIGAGVTFHDDDLIP 60  
Qy 60 FGASDTEREAHVKRFRQALDGMVPMATTNLFTHPVFKAGFTANDRAVRRYALRXTI 119  
Db 61 FGSSDTEREAHVKRFRQALDGMVPMATTNLFTHPVFKAGFTANDRAVRRYALRXTI 120

Qy 120 RNIDLAVELGAKVYVAMGREGAESGAADKDYRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179  
Db 121 RNIDLAVELGAKVYVAMGREGAESGAADKDYRAALDRMKEAFDILLGEYVTSQGYDIRFAI 180  
Qy 180 EPKPNPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMGAGLNFPHGIAQALWAGK 239  
Db 181 EPKPNPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMGAGLNFPHGIAQALWAGK 240  
Qy 240 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLESAGYEGPRHDFKPPRTEDFDGVM 299  
Db 241 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLESAGYEGPRHDFKPPRTEDFDGVM 300  
Qy 300 ASAGCMRNLYLKERAAFRADPEVQALRAALDQALAEPTAADGLQALLADRTAYEDF 359  
Db 301 ASAGCMRNLYLKERAAFRADPEVQALRAALDQALAEPTAADGLQALLADRTAYEDF 360  
Qy 360 DVD-AAARGMAFERLDQALMDHLLGARG 386  
Db 361 DVDAARAGMAFERLDQALMDHLLGARG 388  
RESULT 3  
S28986  
xylose isomerase (EC 5.3.1.5) - Streptomyces olivochromogenes  
C:Species: Streptomyces olivochromogenes  
C:Date: 25-Feb-1994 #sequence\_revision 06-Dec-1996 #text\_change 10-Sep-1999  
C:Accession: S28986  
R:Farber, G.K.; Glasfeld, A.; Tiraby, G.; Ringe, D.; Petsko, G.A.  
Biochemistry 28, 7289-7297, 1989  
A:Title: Crystallographic studies of the mechanism of xylose isomerase.  
A:Reference number: S28986; MUID:90057422  
A:Contents: X-ray crystallography, 3.0 angstroms  
A:Accession: S28986  
A:Molecule type: protein; DNA  
A:Residues: 1-389 <FAR>  
R:Farber, G.; Petsko, G.  
submitted to the Brookhaven Protein Data Bank, February 1989  
A:Reference number: A50637; PDB:3XIA  
A:Contents: annotation: X-ray crystallography, 3.0 angstroms, residues 2-378  
C:Complex: homodimer  
C:Function:  
A:Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopento  
A:Pathway: xylose metabolism  
C:Superfamily: xylose isomerase  
C:Keywords: homodimer; intramolecular oxidoreductase; isomerase; magnesium; xylose me  
F:2-389/Product: xylose isomerase #status experimental <MAT>  
F:54.57/Active site: His, Asp #status experimental  
F:90.183/Binding site: substrate (Thr, Lys) #status experimental  
F:181,217,245,287/Binding site: magnesium (Glu, Glu, Asp, Asp) #status experimental  
F:217,220,255,257/Binding site: magnesium (Glu, His, Asp, Asp)

Query Match 92.3%; Score 1872.5; DB 1; Length 389;  
Best Local Similarity 91.5%; Pred. No. 1e-135;  
Matches 356; Conservative 15; Mismatches 15; Indels 3; Gaps 2;  
Qy 1 MNQYPTPEDRTFFGLTWVGWQGRDPFGDTRPALDPVD-VORLAEIGAYGVTFHDDDLIP 59  
Db 1 MSYQYPTPEDRTFFGLTWVGWQGRDPFGDTRPALDPVDVORLAEIGAGVTFHDDDLIP 60  
Qy 60 FGASDTEREAHVKRFRQALDGMVPMATTNLFTHPVFKAGFTANDRAVRRYALRXTI 119  
Db 61 FGSSDTEREAHVKRFRQALDGMVPMATTNLFTHPVFKAGFTANDRAVRRYALRXTI 120  
Qy 120 RNIDLAVELGAKVYVAMGREGAESGAADKDYRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179  
Db 121 RNIDLAVELGAKVYVAMGREGAESGAADKDYRAALDRMKEAFDILLGEYVTSQGYDIRFAI 180  
Qy 180 EPKPNPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMGAGLNFPHGIAQALWAGK 239  
Db 181 EPKPNPRGDILLPTIGHALAFIERLERPELYGVNPEVGHQMGAGLNFPHGIAQALWAGK 240  
Qy 240 LFHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLESAGYEGPRHDFKPPRTEDFDGVM 299



Db 241 LFHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGYEGPRHLDFKPPRTEDIDGVW 300  
 QY 300 ASAGCMRNLYLKERAAAFRADEVEQALRAARLDOLAEPDADGQALLADRTAYEDF 359  
 Db 301 ASAGCMRNLYLKERAAAFRADEVEQALRAARLDOLAEPDADGQALLADRTAYEDF 360  
 QY 360 DVDAARGMA--FERLDOLAMHLLGARG 386  
 Db 361 DVDAARGMA--FERLDOLAMHLLGARG 389

RESULT 4  
 JC1031  
 xylose isomerase (EC 5.3.1.5) - Streptomyces diastaticus  
 C:Species: Streptomyces diastaticus  
 C:Date: 30-Sep-1993 #sequence\_revision 19-Oct-1995 #text\_change 06-Dec-1996  
 C:Accession: JC1031

R;Wang, Y.Z.; Huang, Z.; Dai, X.H.; Liu, J.; Cui, T.; Niu, L.W.; Wang, C.; Xu, X.  
 Chinese J. Biotechnol. 10, 118-123, 1994  
 A:Title: DNA sequence of xylose isomerase gene from Streptomyces diastaticus No.7 strain  
 A:Reference number: JC1031  
 A:Contents: No.7, M1033  
 A:Accession: JC1031  
 A:Molecule type: DNA  
 A:Residues: 1-388 <WAN>  
 C:Superfamily: xylose isomerase  
 C:Keywords: intramolecular oxidoreductase; isomerase

Query Match 90.2%; Score 1831; DB 2; Length 388;  
 Best Local Similarity 89.7%; Pred. No. 1.5e-132;  
 Matches 348; Conservative 17; Mismatches 21; Indels 2; Gaps 2;

QY 1 MNYQTPEDRFTGLTWGQGRDPFGDTRPALDPVD-VORLAELGAYGVTHDDDLIP 59  
 Db 1 MSQTPEDRFTGLTWGQGRDPFGDTRPALDPVD-VORLAELGAYGVTHDDDLIP 60  
 QY 60 FGASDTEREAHVRRFOALDGMTVPMATNLTFTPVFKAGAFANDRAVRYALRKT 119  
 Db 61 FGATDSERAEHRRFOGLDETKMKVPMATNLTFTPVFKAGAFANDRAVRYALRKT 120  
 QY 120 RNTDLAVELGAKYVAVWGREGAESGAADKVRALDRMKEAFDLGLGYVTSQGYDIRFAI 179  
 Db 121 RNTDLAVELGAKYVAVWGREGAESGAADKVRALDRMKEAFDLGLGYVTSQGYDIRFAI 180  
 QY 180 EKPNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGK 239  
 Db 181 EKPNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGK 240  
 QY 240 LFHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGYEGPRHLDFKPPRTEDIDGVW 299  
 Db 241 LFHIDLNGQSGIKYDQDLRFAGDLRAAFWLVLDLLESAGYEGPRHLDFKPPRTEDIDGVW 300  
 QY 300 ASAGCMRNLYLKERAAAFRADEVEQALRAARLDOLAEPDADGQALLADRTAYEDF 359  
 Db 301 ASAGCMRNLYLKERAAAFRADEVEQALRAARLDOLAEPDADGQALLADRTAYEDF 360  
 QY 360 DVD-AAARGMAFERLDOLAMHLLGARG 386  
 Db 361 DPDAARGMAFERLDOLAMHLLGARG 388

RESULT 5  
 ISSMR  
 xylose isomerase (EC 5.3.1.5) - Streptomyces rochei  
 N:Alternate names: xylose ketol-isomerase  
 C:Species: Streptomyces rochei  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 07-May-1999  
 C:Accession: JN0086  
 R;Kikuchi, T.; Itoh, Y.; Kasumi, T.; Fukazawa, C.  
 Agric. Biol. Chem. 54, 2469-2472, 1990  
 A:Title: Molecular cloning of the xylA gene encoding xylose isomerase from Streptomyces

A:Reference number: JN0086; MUID:91136786  
 A:Accession: JN0086  
 A:Molecule type: DNA  
 A:Residues: 1-394 <XIK>  
 C:Comment: Phe-26, Phe-94, and Trp-136 have been demonstrated to be involved in facill  
 C:Genetics:  
 A:Gene: xylA  
 C:Complex: homotetramer  
 C:Function:  
 A:Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopento  
 A:Pathway: xylose metabolism  
 C:Superfamily: xylose isomerase  
 C:Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; xylose  
 F:2-394/Product: xylose isomerase #status predicted <MAT>  
 F:54/57/Active site: His, Asp #status predicted  
 F:90/182/Binding site: substrate (Thr, Lys) #status predicted  
 F:180/216/244/285/Binding site: magnesium (Glu, Glu, Asp, Asp) #status predicted  
 F:216/219/254/256/Binding site: magnesium (Glu, His, Asp, Asp) #status predicted

Query Match 75.9%; Score 1539.5; DB 1; Length 394;  
 Best Local Similarity 79.3%; Pred. No. 2.7e-110;  
 Matches 306; Conservative 21; Mismatches 24; Indels 35; Gaps 8;

QY 1 MNYQTPEDRFTGLTWGQGRDPFGDTRPALDPVD-VORLAELGAYGVTHDDDLIP 59  
 Db 1 MSQTPEDRFTGLTWGQGRDPFGDTRPALDPVD-VORLAELGAYGVTHDDDLIP 60  
 QY 60 FGASDTEREAHVRRFOALDGMTVPMATNLTFTPVFKAGAFANDRAVRYALRKT 119  
 Db 61 FGSDTERESHIRFRQALDGMTVPMATNLTFTPVFK-DRTANDRDRVAYAVRKT 119  
 QY 120 RNTDLAVELGAKYVAVWGREGAESGAADKVRALDRMKEAFDLGLGYVTSQGYDIRFAI 179  
 Db 120 RNTDLAVELGAKYVAVWGREGAESGAADKVRALDRMKEAFDLGLGYVTSQGYDIRFAI 179  
 QY 180 EKPNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGK 239  
 Db 180 EKPNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGEQAGLNFPHGIAQALWAGK 239  
 QY 240 LFHIDLNGQSGIKYDQDL--REFAGDLRAAFWLVLDLLESAGYEGPRHLDFKPPRTEDID 296  
 Db 240 LFHIDLNGQSGIKYDQCGSRRRPPAGV----LVVDLLESAGYEGPRHLDFKPPRTEDID 295  
 QY 297 GVWASAGCMRNLYLKE-----RAAFRADEVEQALRAARLDOLAEPDAA 343  
 Db 296 GVWASAGCMRNLYLKPSPPSAPTRRCRRASA----PRVWTSWPSRPL-----A 343  
 QY 344 DGLQALLADRTAYEDFVDV-AAARGM 368  
 Db 344 DGLQALLADRTAYEDFVDEAAARGM 369

RESULT 6  
 ISMAXA  
 xylose isomerase (EC 5.3.1.5) - Ampullariella sp.  
 N:Alternate names: xylose ketol-isomerase  
 C:Species: Ampullariella sp.  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 06-Dec-1996  
 C:Accession: A27756  
 R;Saari, G.C.; Kumar, A.A.; Kawasaki, G.H.; Insley, M.Y.; O'Hara, P.J.  
 J. Bacteriol. 169, 612-618, 1987  
 A:Title: Sequence of the Ampullariella sp. strain 3876 gene coding for xylose isomera  
 A:Reference number: A27756; MUID:87109047  
 A:Accession: A27756  
 A:Molecule type: DNA  
 A:Residues: 1-394 <SAA>  
 A:Cross-references: GB:M15050  
 A:Experimental source: strain 3876  
 C:Genetics:  
 A:Start codon: GTG  
 C:Complex: homotetramer  
 C:Function:

A;Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopentose  
A;Pathway: xylose metabolism  
C;Superfamily: xylose isomerase  
C;Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; xylose me  
F:54,57/Active site: His, Asp #status predicted  
F:90,183/Binding site: substrate (Thr, Lys) #status predicted  
F:181,217,245,292/Binding site: magnesium (Glu, Asp, Asp) #status predicted  
F:217,220,255,257/Binding site: magnesium (Glu, His, Asp, Asp) #status predicted

Query Match 66.7%; Score 1353.5; DB 1; Length 394;  
Best Local Similarity 66.0%; Pred. No. 4.3e-96;  
Matches 260; Conservative 46; Mismatches 79; Indels 9; Gaps 4;

```
Qy 1 MNVQTPEDRTFFGLTWVGWGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 59
Db 1 MSLQATPDOKFSGFLWTVGWAQDAFGDTRPVLPDPIEAVHKLAEIGAYGVTFHDDDLVP 60

Qy 60 FGASDTEREAHVYKREFRQALDGMTVPMTNLFTHPVFKAGAFANDRAVRRYALRXTI 119
Db 61 FGADAATRDGIVAGFSKALDETGLVPMVTNLFTHPVFKDGGFTSNDRSVRRYAIRKVL 120

Qy 120 RNIDLAELGAKVYVWAGREGAESAADKDYRAALDRMKEAFDILGEYVTSOGYDIRFAI 179
Db 121 ROMDLGAELGAKTLLVWGREGAEDYSAKDYGAAALDRYREALNLQAQYSEDQGYGLPFAI 180

Qy 180 EPKPNPRGDIILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 239
Db 181 EPKPNPRGDIILLPTAGHAIQVQELERPELFGINPETHGQMSNLNFTQGIQAALWHKK 240

Qy 240 LPHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLLESA-----GWEGPRHDFKPPRTED 294
Db 241 LPHIDLNGHGPKEQDQDLVFGHDLNAPSLVDLLENGPDGPGAYDGGPRHDFKPSRTED 300

Qy 295 IDGVWASAGCNRNVLILKERAAFRADPEVQEAALRAALDQAEPTAAG--LQALLAD 352
Db 301 FDGVWESAKDNIRMYLLKRAKAFRADEPQVQAALAESKVDLRTPTLNPGETYADLLAD 360

Qy 353 RTAYEDFDVDA-AARGMAFERLDQLAMDLHLLGAR 385
Db 361 RSAPEDYDADAVGAKGFGVKNLQNLAIHLLGAR 394
```

RESULT 7  
ISMAXH  
xylose isomerase (EC 5.3.1.5) - Actinoplanes missouriensis  
N:Alternate names: xylose ketol-isomerase  
C:Species: Actinoplanes missouriensis  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 18-Jun-1999  
C:Accession: S05998  
R:Amore, R.; Hollenberg, C.P.  
Nucleic Acids Res. 17, 7515, 1989  
A:Title: Xylose isomerase from Actinoplanes missouriensis: primary structure of the gene  
A:Reference number: S05998; PIDN:90016811  
A:Accession: S05998  
A:Molecule type: DNA  
A:Residues: 1-394 <CDS>  
A:Cross-references: EMBL:X16042; NID:g38870; PIDN:CA34164.1; PID:g580713  
A:Note: the authors translated the initiation codon GTG for residue 1 as Val  
R:Janin, J.  
submitted to the Brookhaven Protein Data Bank, April 1992  
A:Reference number: A51549; PDB:6XIM  
C:Contents: annotation; X-ray crystallography, 2.5 angstroms, with xylose and magnesium,  
R:Rev, F.; Jenkins, J.; Janin, J.; Lasters, I.; Alard, P.; Claessens, M.; Matthysens, G  
Proteins 4, 165-172, 1988  
A:Title: Structural analysis of the 2.8 angstroms model of xylose isomerase from Actinop  
A:Reference number: A58417; MUID:89184498  
C:Contents: annotation; X-ray crystallography, 2.8 angstroms  
C:Genetics:  
A:Start codon: GTG  
C:Function:  
A:Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopentose

A;Pathway: xylose metabolism  
C;Superfamily: xylose isomerase  
C;Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; xylose  
F:54,57/Active site: His, Asp #status experimental  
F:90,183/Binding site: substrate (Thr, Lys) #status experimental  
F:181,217,245,292/Binding site: magnesium (Glu, Asp, Asp) #status experimental  
F:217,220,255,257/Binding site: magnesium (Glu, His, Asp, Asp) #status experimental

Query Match 66.6%; Score 1351.5; DB 1; Length 394;  
Best Local Similarity 65.5%; Pred. No. 6.1e-96;  
Matches 258; Conservative 51; Mismatches 76; Indels 9; Gaps 4;

```
Qy 1 MNVQTPEDRTFFGLTWVGWGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 59
Db 1 MSVQATREDKESFGLWTVGWAQDAFGDTRTALDPVAVHKLAEIGAYGVTFHDDDLVP 60

Qy 60 FGASDTEREAHVYKREFRQALDGMTVPMTNLFTHPVFKAGAFANDRAVRRYALRXTI 119
Db 61 FGSDAQTDRDGIAGFKALDETGLVPMVTNLFTHPVFKDGGFTSNDRSVRRYAIRKVL 120

Qy 120 RNIDLAELGAKVYVWAGREGAESAADKDYRAALDRMKEAFDILGEYVTSOGYDIRFAI 179
Db 121 ROMDLGAELGAKTLLVWGREGAEDYSAKDYSAAALDRYREALNLQAQYSEDQGYGLRFAI 180

Qy 180 EPKPNPRGDIILLPTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 239
Db 181 EPKPNPRGDIILLPTAGHAIQVQELERPELFGINPETHGQMSNLNFTQGIQAALWHKK 240

Qy 240 LPHIDLNGSGIKYDQDLRFAGDGLRAAFWLVLDLLESA-----GWEGPRHDFKPPRTED 294
Db 241 LPHIDLNGHGPKEQDQDLVFGHDLNAPSLVDLLENGPDGAPAYDGGPRHDFKPSRTED 300

Qy 295 IDGVWASAGCNRNVLILKERAAFRADPEVQEAALRAALDQAEPT--AADGLQALLAD 352
Db 301 YDGVWESAKDNIRMYLLKRAKAFRADEPQVQAALAESKVDLRTPTLNPGETYADLLAD 360

Qy 353 RTAYEDFDVDA-AARGMAFERLDQLAMDLHLLGAR 385
Db 361 RSAPEDYDADAVGAKGFGVKNLQNLAIHLLGAR 394
```

RESULT 8  
S16214  
xylose isomerase (EC 5.3.1.5) - Arthrobacter sp.  
N:Alternate names: D-glucose isomerase  
C:Species: Arthrobacter sp.  
C:Date: 21-Nov-1993 #sequence\_revision 06-Dec-1996 #text\_change 18-Jun-1999  
C:Accession: S16214; S16212; S28154  
R:Loviny-Anderson, T.; Shaw, P.C.; Shin, M.K.; Hartley, B.S.  
Biochem. J. 277, 263-271, 1991  
A:Title: D-Xylose (D-glucose) isomerase from Arthrobacter strain N.R.R.L. B3728. Gene  
A:Reference number: S16213; MUID:91307531  
A:Accession: S16214  
A:Molecule type: DNA  
A:Residues: 1-395 <BIO>  
A:Cross-references: EMBL:X59466; NID:g39058; PIDN:CAA42073.1; PID:g39059  
A:Experimental source: strain NRRL B3728  
R:Smith, C.A.; Rangarajan, M.; Hartley, B.S.  
Biochem. J. 277, 255-261, 1991  
A:Title: D-Xylose (D-glucose) isomerase from Arthrobacter strain N.R.R.L. B3728. Puri  
A:Reference number: S16212; MUID:91307530  
A:Accession: S16212  
A:Molecule type: protein  
A:Residues: 2-21 <SMI>  
A:Experimental source: strain NRRL B3728  
A:Note: other peptide sequences were also determined  
R:Siddiqui, K.S.; Rangarajan, M.; Hartley, B.S.; Kitmitto, A.; Panico, M.; Blench, I.  
Biochem. J. 289, 201-208, 1993  
A:Title: Arthrobacter D-xylose isomerase: partial proteolysis with thermolysin.  
A:Reference number: S28154; MUID:93143669  
A:Accession: S28154  
A:Molecule type: protein

A:Residues: 337-348 <STD>  
R:Henrick, K.; Collier, C.A.; Blow, D.M.  
submitted to the Brookhaven Protein Data Bank, July 1989  
A:Reference number: A50702; PDB:4XIA  
A:Contents: annotation: X-ray crystallography, 2.3 angstroms, with D-sorbitol, residues  
R:Henrick, K.; Collier, C.A.; Blow, D.M.  
J. Mol. Biol. 208, 129-157, 1989  
A:Title: Structures of D-xylose isomerase from *Arthrobacter* strain B3728 containing the  
A:Reference number: A58418; MUID:89362453  
A:Contents: annotation: X-ray crystallography, 2.3 angstroms  
R:Collier, C.A.; Henrick, K.; Blow, D.M.  
submitted to the Brookhaven Protein Data Bank, October 1991  
A:Reference number: A51428; PDB:1XLC  
A:Contents: annotation: X-ray crystallography, 2.5 angstroms, with D-xylose/xyulose and  
R:Collier, C.A.; Henrick, K.; Blow, D.M.  
J. Mol. Biol. 212, 211-235, 1990  
A:Title: Mechanism for aldose-ketose interconversion by D-xylose isomerase involving ring  
A:Reference number: A58419; MUID:90704542  
A:Contents: annotation: X-ray crystallography, 2.5 angstroms  
C:Genetics:  
A:Gene: xylA  
C:Complex: homotetramer  
C:Function:  
A:Description: catalyzes isomerization of the aldopentose D-xylose to the 2-ketopentose  
A:Pathway: xylose metabolism  
C:Superfamily: xylose isomerase  
C:Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; xylose me  
F:2-395/Product: xylose isomerase #status experimental <MAT>  
F:54.57/Active site: His, Asp #status experimental  
F:90.183/Binding site: substrate (Thr, Lys) #status experimental  
F:181,217,245,293/Binding site: magnesium (Glu, Glu, Asp, Asp) #status experimental  
F:217,220,255,257/Binding site: magnesium (Glu, His, Asp, Asp) #status experimental

Query Match 63.2%; Score 1282; DB 1; Length 395;  
Best Local Similarity 62.3%; Pred. No. 1.2e-90;  
Matches 246; Conservative 53; Mismatches 86; Indels 10; Gaps 4;

QY 1 MNYQPTPEDRFTGLTWVGWGRDPFGDTRPALDPVD-VORLAELGAYGVTFHDDDLIP 59  
DB 1 MSVQPTPDHFTGLTWGTVGADPGVATRNKLDPEAVHKLAEALGAYGITHDNDLIP 60

QY 60 FGASDTEREAHVKRFRQALDAGMTVPMTATNLTFTHPVKAGAFANDRAVRYALRKTI 119  
DB 1 FDATEAREKILGDFNQALDKDTGLKVPMTNLTFSHPVFKDGGFTSNDRSIRFALAKVL 120

QY 120 RNIDLAELGAKYVYVWVGREGAESGAADVRAALDRMKEAFDILLGEYVTSQGYDIRFAI 179  
DB 121 HNIDLAEMGAETFFVWVGREGSEYDGSKDAAALDRMGVDTAAGYIKDKGINLRUAL 180

QY 180 EKPNEPRGDIILPTTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGK 239  
DB 181 EKPNEPRGDIPLTVGHGLAFIEQLHEHGDIVGLNPEVGHQMAGLNFTHGIAQALWAEK 240

QY 240 LEHIDLUNGSGIKYDQDLRFAGAGDLRAAFWLVDLSEA-----GWEGPRHDFKPPRTE 293  
DB 241 LEHIDLUNGQGIKDYDQLFVGHGDLTSAPFTVDLLENGPFGPKYKTPRHFYKPSRTD 300

QY 294 DIDGVWASAGCRNRYLLIKERAAAFRAADPEVQALRAALDOLAEP--AADGLQALLA 351  
DB 301 GYDGVWDSAKANNMYSLLIKERAAAFRAADPEVQALRAALDOLAEP--AADGLQALLA 360

QY 352 DRTAVEDFDVDA--RGMAFERLDOLAMHLLGAR 385  
DB 361 DSNASFGAFDAEAAERNAFIRLNQALAEHLGSR 395

RESULT 9  
A39404  
xylose isomerase (EC 5.3.1.5) - *Thermus aquaticus*  
N:Alternate names: xylose ketol-isomerase  
C:Species: *Thermus aquaticus*  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 31-Dec-1993

C:Accession: A39404  
R:Dekker, K.; Yamagata, H.; Sakaguchi, K.; Uda, S.  
J. Bacteriol. 173, 3078-3083, 1991  
A:Title: Xylose (glucose) isomerase gene from the thermophile *Thermus thermophilus*: c  
A:Reference number: A39404; MUID:91216978  
A:Accession: A39404  
A:Molecule type: DNA  
A:Residues: 1-387 <DEK>  
C:Superfamily: xylose isomerase  
C:Keywords: homotetramer; intramolecular oxidoreductase; isomerase; magnesium; pentos

Query Match 55.5%; Score 1127; DB 2; Length 387;  
Best Local Similarity 58.5%; Pred. No. 8.3e-79;  
Matches 228; Conservative 43; Mismatches 109; Indels 10; Gaps 3;

QY 3 YOPTPEDRFTGLTWVGWGRDPFGDTRPALDPV-DVORLAELGAYGVTFHDDDLIPFG 61  
DB 2 YEPKPEHRTFTGLTWVGWGRDPFGDTRPALDPV-DVORLAELGAYGVTFHDDDLIPFG 61

QY 62 ASDTEREAHVKRFRQALDAGMTVPMTATNLTFTHPVKAGAFANDRAVRYALRKTI 121  
DB 62 TPOQERDQIVRRFKALDETGLKVPMTANLFSDDPAFKDGAFTSPDPWVRYALRKSL 121

QY 122 IDLAELGAKYVYVWVGREGAESGAADVRAALDRMKEAFDILLGEYVTSQGYDIRFAI 181  
DB 122 MDLGAELGAEIYVWVGREGAEVATGKARKVMDWYREALNFMAAYAEQGYGFAL 181

QY 182 KNEPRGDIILPTTIGHALAFIERLERPELYGVNPEVGHQMAGLNFPHGIAQALWAGKLF 241  
DB 182 KNEPRGDIYFATVGSMLAFITLDRPERFGLNPEFAHETMAGLNFVHVAQALDAGKLF 241

QY 242 HIDLUNGSGIKYDQDLRFAGAGDLRAAFWLVDLLESAAGWEGPRHDFKPPRTEIDIGVWAS 301  
DB 242 HIDLNDORMSRFDQDLRFSGENLKAAFFLVLDLLESSGYQGRHFDHALRTEDEEGWAF 301

QY 302 AGCMRNYLLIKERAAAFRAADPEVQALRAALDQ-----LAEPTAADGLQALLADRTAY 356  
DB 302 ARGCMRTYLLIKERAAAFRAADPEVQALRAALDQ-----LAEPTAADGLQALLADRTAY 356

QY 357 EDFDVAAGMAFERLDOLAMHLLGAR 386  
DB 362 E-----AKRRRGYALERLDOLAVEYLLGVRG 387

RESULT 10  
S18561  
xylose isomerase (EC 5.3.1.5) - *Lactobacillus pentosus*  
C:Species: *Lactobacillus pentosus*  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 16-Jul-1999  
C:Accession: S18561  
R:Lokman, B.C.; van Santen, P.; Verdoes, J.C.; Kruese, J.; Leer, R.J.; Posno, M.; Pou  
Mol. Gen. Genet. 230, 161-169, 1991  
A:Title: Organization and characterization of three genes involved in D-xylose catabo  
A:Reference number: S18560; MUID:92079891  
A:Accession: S18561  
A:Molecule type: DNA  
A:Residues: 1-449 <LOK>  
A:Cross-references: EMBL:N57384; NID:g149604; PIDN:AAA25258.1; PID:g149606  
C:Genetics:  
A:Gene: xylA  
C:Superfamily: xylose isomerase  
C:Keywords: intramolecular oxidoreductase; isomerase; magnesium; xylose metabolism  
F:103.106/Active site: His, Asp #status predicted  
F:143,236/Binding site: substrate (Thr, Lys) #status predicted  
F:234,270,298,342/Binding site: magnesium (Glu, Glu, Asp, Asp) #status predicted

Query Match 16.4%; Score 333.5; DB 2; Length 449;  
Best Local Similarity 29.8%; Pred. No. 3.9e-18;  
Matches 122; Conservative 61; Mismatches 150; Indels 77; Gaps 18;





Matches 101; Conservative 53; Mismatches 149; Indels 35; Gaps 11;

```
Qy 10 RTFTGLW-TVGMQGRDRPGDAT--RP---ALDPVDVOR-----LAEIGAYGVTFHD 54
Db 43 RFSIAYHTFTADGTDQFGKATMORPNHYTDPMDIAKARVEAAFEFEDKINAFYFCFHD 102
Qy 55 DDLIPFGASDTEREAH-----VKRFQALDATGMTVPMATNLTHTPVFKAGAFANDR 108
Db 103 RDIAPEG--DPLRETNKNLDTIVAMINDYLATSKTKVLWGTAFLFSNPRFVHGASTSCNA 160
Qy 109 AVRRYALRKTIRNIDLAVELGAKVYVWGGREGAESGAADVRAALDRMKEAFDILLGEYV 168
Db 161 DVFAYSAAQVKALEITKELGGENYVFWGGREGYETLNTDMEFELDNFARFLHMAVDYA 220
Qy 169 TSQGVDIRFAIEPKPNEPRGDIILLPTIGHALAFIERLERPELYGVNPEVGEHQMAGLNFP 228
Db 221 KEIGFEGQFLIEPAKPEKTKHQYDFDVANVLAFLKRYDLDKRYFKVNIENHATLAFHDFQ 280
Qy 229 HGIAQALWAGKLFHIDLN-GOSGIKYDQDLRFAGDLRAAFWLYDVLLESAGWE-GPRHFD 286
Db 281 HELRYARINGVLSIDANTGDMLLGWDTD-QFPTDIRMTTLAMVEVIKMGGFDKGGLNFD 339
Qy 287 FKPPRT-----ED-----IDGVWASAAGCMRNYLILKER 315
Db 340 AKVRRASPEDEFLGHIAGMDAFAGKFKVAYKLVKDR 377
```

Search completed: March 22, 2000, 23:44:40  
Job time: 776 sec